GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 15, 2004, 21:25:56 ; Search time 6853 Seconds
(without alignments)
6830.655 Million cell updates/sec

Title: US-10-759-277-3
Perfect score: 1080
Sequence: 1 atgccgctggagctggagct.....gcaaaaaaatcaagtcttaa 1080

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1080	100.0	1080	6	BD270235	BD270235 Human gly
2	1080	100.0	1080	9	AF257466	AF257466 Homo sapi
3	1076.8	99.7	1181	9	BC019315	BC019315 Homo sapi
4	1076.8	99.7	1229	9	BC000008	BC000008 Homo sapi
5	1076.8	99.7	1230	6	AR339544	AR339544 Sequence
6	1075.2	99.6	1170	6	AX877224	AX877224 Sequence
7	1075.2	99.6	1170	6	BD156549	BD156549 Primer fo
8	1075.2	99.6	1170	9	AK001659	AK001659 Homo sapi
9	883.2	81.8	1247	10	BC003307	BC003307 Mus muscu
10	883.2	81.8	1916	10	BC057977	BC057977 Mus muscu
11	881.6	81.6	1174	10	AB041263	AB041263 Mus muscu
12	732.4	67.8	3026	9	AK091184	AK091184 Homo sapi
13	710.2	65.8	813	6	AX868777	AX868777 Sequence
14	710.2	65.8	813	6	BD148839	BD148839 Primer fo
c 15	483.8	44.8	514	6	AX873806	AX873806 Sequence
c 16	483.8	44.8	514	6	BD153868	BD153868 Primer fo
17	456.6	42.3	237317	2	AC097420	AC097420 Rattus no
c 18	418	38.7	464	6	AX408867	AX408867 Sequence
c 19	386.8	35.8	553	6	AX557002	AX557002 Sequence
20	338.6	31.4	228917	2	AC106079	AC106079 Rattus no
c 21	338.6	31.4	229321	2	AC094185	AC094185 Rattus no
22	280.6	26.0	404	6	AX201885	AX201885 Sequence
c 23	269.8	25.0	2044	9	AK126144	AK126144 Homo sapi
24	269.8	25.0	201300	9	AL137073	AL137073 Human DNA
25	241.6	22.4	227816	10	AL683884	AL683884 Mouse DNA
26	241.4	22.4	244134	2	AC097073	AC097073 Rattus no
27	221	20.5	1194	3	AF397531	AF397531 Drosophil
c 28	210.4	19.5	1999	9	AK123739	AK123739 Homo sapi
29	199.6	18.5	197522	2	AC137022	AC137022 Rattus no
c 30	195	18.1	182381	2	AC122081	AC122081 Rattus no
c 31	195	18.1	191841	2	AC112827	AC112827 Rattus no
c 32	195	18.1	266461	2	AC120718	AC120718 Rattus no
33	145.6	13.5	292100	1	SCO939121	AL939121 Streptomy
34	144.4	13.4	11237	1	AE012850	AE012850 Chlorobiu
c 35	135	12.5	302300	1	AP005034	AP005034 Streptomy
36	126.4	11.7	298050	1	BX321861	BX321861 Nitrosomo
c 37	125	11.6	349652	1	BX569690	BX569690 Synechoco
38	119.8	11.1	15267	3	AB055099	AB055099 Drosophil
39	119.8	11.1	22990	3	AB055101	AB055101 Drosophil
40	119.8	11.1	23007	3	AB055100	AB055100 Drosophil
41	119.8	11.1	23244	3	AB055098	AB055098 Drosophil
c 42	119.8	11.1	61204	2	AC017132	AC017132 Drosophil
43	119.8	11.1	161601	3	AC007594	AC007594 Drosophil
44	119.8	11.1	225655	3	AE003695	AE003695 Drosophil
c 45	111.8	10.4	86896	1	RCU57682	U57682 Rhodobacter

ALIGNMENTS

RESULT 1
BD270235
LOCUS BD270235 1080 bp DNA linear PAT 17-JUL-2003
DEFINITION Human glycosylation enzymes.
ACCESSION BD270235
VERSION BD270235.1 GI:33080003
KEYWORDS JP 2002537796-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1080)
AUTHORS Coleman,T.A. and Betenbaugh,M.J.
TITLE Human glycosylation enzymes
JOURNAL Patent: JP 2002537796-A 2 12-NOV-2002;

```
            HUMAN GENOME SCIENCES INC,JOHNS HOPKINS UNIVERSITY
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002537796-A/2
            PD   12-NOV-2002
            PF   01-MAR-2000 JP 2000602748
            PR   02-MAR-1999 US    60/122409
            PI   TIMOTHY A COLEMAN,MICHAEL J BETENBAUGH
            PC   C12N15/09,A61K31/711,A61K38/43,A61K48/00,A61P1/02,A61P3/00, PC
              A61P7/00,
            PC   A61P7/02,A61P7/06,A61P11/00,A61P13/12,A61P15/00,A61P17/02, PC
             A61P19/02,
            PC   A61P19/10,A61P21/04,A61P25/00,A61P25/16,A61P25/24,A61P25/28,
            PC   A61P29/00,
            PC   A61P31/18,A61P35/00,A61P37/00,A61P43/00,C07K16/40,C12N1/15, PC
              C12N1/19,
            PC   C12N1/21,C12N5/10,C12N9/00,G01N33/15,G01N33/50,G01N33/53, PC
            C12N15/00,
            PC   C12N5/00,A61K37/48
            CC   Human glycosylation enzymes
            FH   Key             Location/Qualifiers
            FT   CDS             (1). .(1077).
FEATURES              Location/Qualifiers
     source           1. .1080
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 1080;  DB 6;  Length 1080;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-256;
  Matches 1080;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080

RESULT 2
AF257466
LOCUS       AF257466               1080 bp    mRNA    linear   PRI 11-JUN-2000
DEFINITION  Homo sapiens N-acetylneuraminic acid phosphate synthase mRNA,
            complete cds.
ACCESSION   AF257466
VERSION     AF257466.1  GI:8453155
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1080)
  AUTHORS   Lawrence,S.M., Huddleston,K.A., Pitts,L.R., Nguyen,N., Lee,Y.C.,
            Vann,W.F., Coleman,T.A. and Betenbaugh,M.J.
  TITLE     Cloning and expression of the human N-acetylneuraminic acid
            phosphate synthase gene with 2-keto-3-deoxy-D-glycero-
            D-galacto-nononic acid biosynthetic ability
  JOURNAL   J. Biol. Chem. 275 (23), 17869-17877 (2000)
  MEDLINE   20298869
   PUBMED   10749855
REFERENCE   2  (bases 1 to 1080)
  AUTHORS   Lawrence,S.M., Huddleston,K.A., Pitts,L.R., Nguyen,N., Lee,Y.C.,
            Vann,W.F., Coleman,T.A. and Betenbaugh,M.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-APR-2000) Chemical Engineering, Johns Hopkins
            University, 3400 N. Charles St., Baltimore, MD 21218, USA
FEATURES              Location/Qualifiers
     source           1. .1080
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /tissue_type="liver"
     CDS              1. .1080
                      /note="2-keto-3-deoxy-D-glycero-D-galacto-nononic acid
                      biosynthetic ability"
                      /codon_start=1
                      /product="N-acetylneuraminic acid phosphate synthase"
                      /protein_id="AAF75261.1"
                      /db_xref="GI:8453156"
                      /translation="MPLELELCPGRWVGGQHPCFIIAEIGQNHQGDLDVAKRMIRMAK
```

```
                     ECGADCAKFQKSELEFKFNRKALERPYTSKHSWGKTYGEHKRHLEFSHDQYRELQRYA
                     EEVGIFFTASGMDEMAVEFLHELNVPFFKVGSGDTNNFPYLEKTAKKGRPMVISSGMQ
                     SMDTMKQVYQIVKPLNPNFCFLQCTSAYPLQPEDVNLRVISEYQKLFPDIPIGYSGHE
                     TGIAISVAAVALGAKVLERHITLDKTWKGSDHSASLEPGELAELVRSVRLVERALGSP
                     TKQLLPCEMACNEKLGKSVVAKVKIPEGTILTMDMLTVKVGEPKAYPPEDIFNLVGKK
                     VLVTVEEDDTIMEELVDNHGKKIKS"
ORIGIN

  Query Match          100.0%;  Score 1080;  DB 9;  Length 1080;
  Best Local Similarity  100.0%;  Pred. No. 2.2e-256;
  Matches 1080;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080


RESULT 3
BC019315
LOCUS       BC019315                1181 bp    mRNA    linear   PRI 03-OCT-2003
DEFINITION  Homo sapiens N-acetylneuraminic acid synthase (sialic acid
            synthase), mRNA (cDNA clone MGC:4339 IMAGE:2821389), complete cds.
ACCESSION   BC019315
VERSION     BC019315.1  GI:17939511
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1181)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1181)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-DEC-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.
```

```
          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
          Series: IRAL Plate: 9 Row: d Column: 7
          This clone was selected for full length sequencing because it
          passed the following selection criteria: matched mRNA gi: 12056472.
FEATURES             Location/Qualifiers
     source          1. .1181
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:4339 IMAGE:2821389"
                     /tissue_type="Lung, small cell carcinoma"
                     /clone_lib="NIH_MGC_7"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene            1. .1181
                     /gene="NANS"
                     /note="synonym: SAS"
                     /db_xref="LocusID:54187"
                     /db_xref="MIM:605202"
     CDS             61. .1140
                     /codon_start=1
                     /product="N-acetylneuraminic acid phosphate synthase"
                     /protein_id="AAH19315.1"
                     /db_xref="GI:17939512"
                     /db_xref="LocusID:54187"
                     /translation="MPLELELCPGRWVGGQHPCFIIAEIGQNHQGDLDVAKRMIRMAK
                     ECGADCAKFQKSELEFKFNRKALDRPYTSKHSWGKTYGEHKRHLEFSHDQYRELQRYA
                     EEVGIFFTASGMDEMAVEFLHELNVPFFKVGSGDTNNFPYLEKTAKKGRPMVISSGMQ
                     SMDTMKQVYQIVKPLNPNFCFLQCTSAYPLQPEDVNLRVISEYQKLFPDIPIGYSGHE
                     TGIAISVAAVALGAKVLERHITLDKTWKGSDHSASLEPGELAELVRSVRLVERALGSP
                     TKQLLPCEMACNEKLGKSVVAKVKIPEGTILTMDMLTVKVGEPKGYPPEDIFNLVGKK
                     VLVTVEEDDTIMEELVDNHGKKIKS"
     misc_feature    82. .1095
                     /note="SpsE; Region: Sialic acid synthase [Cell envelope
                     biogenesis, outer membrane]"
                     /db_xref="CDD:COG2089"
ORIGIN

  Query Match             99.7%;  Score 1076.8;  DB 9;  Length 1181;
  Best Local Similarity   99.8%;  Pred. No. 1.4e-255;
  Matches 1078;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 120

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 180

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 240

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        241 AAGTTTAATCGGAAAGCCTTGGACAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 300

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 360

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 420

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 480

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 540

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 600

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 660

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 720

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 780

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 840

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 900

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 960

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 1020

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1080

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1140

RESULT 4
BC000008
LOCUS       BC000008                1229 bp    mRNA    linear   PRI 04-OCT-2003
DEFINITION  Homo sapiens N-acetylneuraminic acid synthase (sialic acid
            synthase), mRNA (cDNA clone MGC:934 IMAGE:3505324), complete cds.
ACCESSION   BC000008
VERSION     BC000008.2  GI:33875092
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1229)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
```

```
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1229)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-2000) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Aug 19, 2003 this sequence version replaced gi:12652538.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
            Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 7 Row: f Column: 1
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 12056472.
FEATURES             Location/Qualifiers
     source          1. .1229
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:934 IMAGE:3505324"
                     /tissue_type="Placenta, choriocarcinoma"
                     /clone_lib="NIH_MGC_21"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene            1. .1229
                     /gene="NANS"
                     /note="synonym: SAS"
                     /db_xref="LocusID:54187"
                     /db_xref="MIM:605202"
     CDS             79. .1158
                     /codon_start=1
                     /product="N-acetylneuraminic acid phosphate synthase"
                     /protein_id="AAH00008.1"
                     /db_xref="GI:12652539"
                     /db_xref="LocusID:54187"
                     /translation="MPLELELCPGRWVGGQHPCFIIAEIGQNHQGDLDVAKRMIRMAK
                     ECGADCAKFQKSELEFKFNRKALERPYTSKHSWGKTYGEHKRHLEFSHDQYRELQRYA
                     EEVGIFFTASGMDEMAVEFLHELNVPFFKVGSGDTNNFPYLEKTAKKGRPMVISSGMQ
                     SMDTMKQVYQIVKPLNPNFCFLQCTSAYPLQPEDVNLRVISEYQKLFPDIPIGYSGHE
                     TGIAISVAAVALGAKVLERHITLDKTWKGSDHSASLEPGELAELVRSVRLVERALGSP
                     TKQLLPCEMACNEKLGKSVVAKVKIPEGTILTMDMLTVKVGEPKGYPPEDIFNLVGKK
                     VLVTVEEDDTIMEELVDNHGKKIKS"
     misc_feature    100. .1113
                     /note="SpsE; Region: Sialic acid synthase [Cell envelope
                     biogenesis, outer membrane]"
                     /db_xref="CDD:COG2089"
ORIGIN

  Query Match             99.7%;  Score 1076.8;  DB 9;  Length 1229;
  Best Local Similarity   99.8%;  Pred. No. 1.4e-255;
  Matches 1078;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 138

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        139 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGATGTAGCCAAGCGCATGATC 198

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 258

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 318

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 378

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 438

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 498

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 558

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 618

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 678

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 738

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 798

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        799 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 858

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        859 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 918

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        919 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 978

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        979 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 1038

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1039 GGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1098

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1099 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1158

RESULT 5
AR339544
LOCUS       AR339544                1230 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 1035 from patent US 6569662.
```

```
ACCESSION   AR339544
VERSION     AR339544.1  GI:33726401
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1230)
  AUTHORS   Tang,Y.T., Zhou,P. and Drmanac,R.T.
  TITLE     Nucleic acids and polypeptides
  JOURNAL   Patent: US 6569662-A 1035 27-MAY-2003;
FEATURES             Location/Qualifiers
     source          1. .1230
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             99.7%;  Score 1076.8;  DB 6;  Length 1230;
  Best Local Similarity   99.8%;  Pred. No. 1.4e-255;
  Matches 1078;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 178

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 238

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 298

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        299 AAGTTTAATCGGAAAGCCTTGGACAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 358

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 418

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 478

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 538

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 598

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 658

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 718

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 778

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 838

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 898

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 958

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 1018

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 1078

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 GGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1138

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1198


RESULT 6
AX877224
LOCUS       AX877224               1170 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 12129 from Patent EP1074617.
ACCESSION   AX877224
VERSION     AX877224.1  GI:40031960
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 12129 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .1170
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             71. .1150
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAE89627.1"
                     /db_xref="GI:40031961"
                     /translation="MPLELELCPGRWVGGQHPCFIIAEIGQNHQGDLDVAKRMIRMAK
                     ECGADCAKFQKSELEFKFNRKALDRPYTSKHSWGKTYGEHKRHLEFSHDQYRELQRYA
                     EEVGIFFTASGMDEMAVEFLHELNVPFFKVGSGDTNNFPYLEKTAKKGRPMVISSGMQ
                     SMDTMKQVYQIVKPLNPNFCFLQCTSAYPLQPEDVNLRVISEYQKLFPDIPIGYSGHE
                     TGIAISVAAVALGTKVLERHITLDKTWKGSDHSASLEPGELAELVRSVRLVERALGSP
                     TKQLLPCEMACNEKLGKSVVAKVKIPEGTILTMDMLTVKVGEPKGYPPEDIFNLVGKK
                     VLVTVEEDDTIMEELVDNHGKKIKS"
ORIGIN

  Query Match             99.6%;  Score 1075.2;  DB 6;  Length 1170;
  Best Local Similarity   99.7%;  Pred. No. 3.4e-255;
  Matches 1077;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 130

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 190

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 250

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
```

```
Db        251 AAGTTTAATCGGAAAGCCTTGGACAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 310
Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
Db        311 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 370
Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
Db        371 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 430
Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
Db        431 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 490
Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
Db        491 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 550
Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
Db        551 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 610
Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
Db        611 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 670
Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
Db        671 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 730
Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
Db        731 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGACCAAGGTGTTGGAACGTCACATAACT 790
Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
Db        791 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 850
Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
Db        851 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 910
Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
Db        911 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 970
Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
Db        971 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 1030
Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
Db       1031 GGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1090
Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
Db       1091 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1150
```

```
RESULT 7
BD156549
LOCUS       BD156549                1170 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD156549
VERSION     BD156549.1  GI:27862307
KEYWORDS    JP 2002191363-A/11392.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1170)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 11392 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/11392
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                      Location/Qualifiers
            FT   CDS             (71). .(1147).
FEATURES                Location/Qualifiers
     source             1. .1170
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
ORIGIN
```

```
  Query Match             99.6%;  Score 1075.2;  DB 6;  Length 1170;
  Best Local Similarity   99.7%;  Pred. No. 3.4e-255;
  Matches 1077;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
Db         71 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 130
Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
Db        131 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 190
Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
Db        191 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 250
Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
Db        251 AAGTTTAATCGGAAAGCCTTGGACAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 310
Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
Db        311 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 370
Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
Db        371 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 430
Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
Db        431 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 490
Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
Db        491 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 550
Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
Db        551 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 610
Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
Db        611 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 670
Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
Db        671 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 730
```

```
Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        731 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGACCAAGGTGTTGGAACGTCACATAACT 790

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        791 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 850

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        851 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 910

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        911 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 970

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        971 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 1030

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1031 GGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1090

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1091 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1150
```

RESULT 8
AK001659
LOCUS AK001659 1170 bp mRNA linear PRI 01-AUG-2002
DEFINITION Homo sapiens cDNA FLJ10797 fis, clone NT2RP4000657, weakly similar to SPORE COAT POLYSACCHARIDE BIOSYNTHESIS PROTEIN SPSE.
ACCESSION AK001659
VERSION AK001659.1 GI:7023052
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuho,Y. and Sasaki,N.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1170)
AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
FEATURES Location/Qualifiers
source 1. .1170
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP4000657"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP4"
/note="cloning vector: pME18SFL3~mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction."
CDS 71. .1150
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91818.1"
/db_xref="GI:7023053"
/translation="MPLELELCPGRWVGGQHPCFIIAEIGQNHQGDLDVAKRMIRMAK
ECGADCAKFQKSELEFKFNRKALDRPYTSKHSWGKTYGEHKRHLEFSHDQYRELQRYA
EEVGIFFTASGMDEMAVEFLHELNVPFFKVGSGDTNNFPYLEKTAKKGRPMVISSGMQ
SMDTMKQVYQIVKPLNPNFCFLQCTSAYPLQPEDVNLRVISEYQKLFPDIPIGYSGHE
TGIAISVAAVALGTKVLERHITLDKTWKGSDHSASLEPGELAELVRSVRLVERALGSP
TKQLLPCEMACNEKLGKSVVAKVKIPEGTILTMDMLTVKVGEPKGYPPEDIFNLVGKK
VLVTVEEDDTIMEELVDNHGKKIKS"
ORIGIN

Query Match 99.6%; Score 1075.2; DB 9; Length 1170;
Best Local Similarity 99.7%; Pred. No. 3.4e-255;
Matches 1077; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 130

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 190

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 250

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        251 AAGTTTAATCGGAAAGCCTTGGACAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 310

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 370

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 430

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 490

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        491 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 550

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 610

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 670

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        671 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 730

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        731 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGACCAAGGTGTTGGAACGTCACATAACT 790

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        791 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 850

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        851 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 910

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        911 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 970

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        971 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 1030

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1031 GGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1090

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1091 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1150
```

RESULT 9
BC003307

```
LOCUS       BC003307               1247 bp    mRNA    linear   ROD 03-OCT-2003
DEFINITION  Mus musculus N-acetylneuraminic acid synthase (sialic acid
            synthase), mRNA (cDNA clone MGC:6982 IMAGE:3154822), complete cds.
ACCESSION   BC003307
VERSION     BC003307.1  GI:13097041
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1247)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1247)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
            A.N., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 6 Row: i Column: 3
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 16716466.
FEATURES             Location/Qualifiers
     source          1. .1247
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="129,C57BL/6J,FVB/N"
                     /db_xref="taxon:10090"
                     /clone="MGC:6982 IMAGE:3154822"
                     /tissue_type="Mammary tumor. Brca1-/fl; MMTV-Cre model. 10
                     months old, gross tissue."
                     /clone_lib="NCI_CGAP_Mam3"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .1247
                     /gene="Nans"
                     /db_xref="LocusID:94181"
                     /db_xref="MGI:2149820"
     CDS             124. .1203
                     /codon_start=1
                     /product="N-acetylneuraminic acid synthase (sialic acid
                     synthase)"
                     /protein_id="AAH03307.1"
                     /db_xref="GI:13097042"
                     /db_xref="LocusID:94181"
                     /translation="MPLELELCPGRWVGGKHPCFIIAEIGQNHQGDIDVAKRMIRTAK
                     ECGADCAKFQKSELEFKFNRKALERPYTSKHSWGKTYGEHKRHLEFSHDQYKELQSYA
                     QEIGIFFTASGMDEMAVEFLHELNVPFFKVGSGDTNNFPYLEKTAKKGRPMVISSGMQ
                     SMDTMKQVYQIVKPLNPNFCFLQCTSAYPLQPEDANLRVISEYQKLFPDIPIGYSGHE
                     TGIAISVAAVALGAKVLERHITLDKTWKGSDHSASLEPGELAELVRSVRLVERALGSP
                     TKQLLPCEMACNEKLGKSVVAKVKIPAGTTLTLDMLTVKVGEPKGYPPEDIFNLAGKK
                     VLVTIEEDDTVMEESVESHSKKIKA"
     misc_feature    145. .1176
                     /note="SpsE; Region: Sialic acid synthase [Cell envelope
                     biogenesis, outer membrane]"
                     /db_xref="CDD:COG2089"
ORIGIN
```

```
  Query Match             81.8%;  Score 883.2;  DB 10;  Length 1247;
  Best Local Similarity   88.6%;  Pred. No. 1.2e-207;
  Matches  957;  Conservative    0;  Mismatches  123;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              |||||||||| |||||||||||||||||||||||||||||| ||  | ||||||||||||
Db        124 ATGCCGCTGGAACTGGAGCTGTGTCCCGGGCGCTGGGTGGGTGGAAAGCACCCGTGCTTC 183

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||| || |||||||||||||||||||| || ||| | || || ||||||||||||||||
Db        184 ATCATCGCGGAGATCGGCCAGAACCACCAAGGAGACATAGATGTGGCCAAGCGCATGATC 243

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||  |||||||||||||||||| || || || ||||| |||||||| ||| | || |||
Db        244 CGCACTGCCAAGGAGTGTGGGGCCGACTGCGCTAAGTTTCAGAAGAGCGAGTTGGAGTTC 303

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||| ||||| ||| |||||||| ||||| || |||||||||||| ||||||||||||
Db        304 AAGTTTAACCGGAAGGCCCTGGAGAGACCATATACTTCGAAGCATTCATGGGGGAAGACG 363

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              || |||||||||| || |||||||| |||||||| |||||||||| ||||||||||||
Db        364 TATGGGGAGCACAAGCGGCATCTGGAATTCAGCCACGACCAGTACAAGGAGCTGCAGAGC 423

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              || ||  ||||| | || ||||||||||||||||||||||||| ||||||||||||||
Db        424 TATGCGCAGGAGATCGGCATCTTCTTCACTGCCTCTGGCATGGACGAGATGGCAGTTGAG 483
```

```
Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              || ||||| |||||||||||||| |||||||||||||||||||| |||||||| ||||||
Db        484 TTTCTGCACGAACTGAATGTTCCCTTTTTCAAAGTTGGATCTGGGGACACTAACAATTTT 543

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              || || |||||||||||||||||| |||||||| || |||||||||||||| |||||||||
Db        544 CCCTACCTGGAAAAGACAGCCAAGAAAGGTCGTCCTATGGTGATCTCCAGCGGGATGCAG 603

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              |||||||||||||||||||||||| |||||||||||||||| || || ||||||||||||
Db        604 TCAATGGACACCATGAAGCAAGTCTATCAGATCGTGAAGCCGCTGAATCCCAACTTCTGC 663

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||| | || || |||||||| ||||| || ||||| |||| | |||||||| ||||||
Db        664 TTCCTCCAATGCACCAGCGCGTACCCACTACAGCCCGAGGATGCCAACCTGCGCGTCATC 723

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              |||||||| |||||||||||||| |||||||||| |||||||| |||| || || |||
Db        724 TCGGAATACCAGAAGCTCTTTCCCGACATTCCCATCGGGTATTCCGGGCACGAGACGGGC 783

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              || || |||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        784 ATCGCCATATCTGTGGCCGCCGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACG 843

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              |||||||||||||||||||||||||||||| |||||||||||||||| || |||||
Db        844 TTGGACAAGACCTGGAAGGGGAGTGACCACTCAGCCTCGCTGGAGCCTGGGGAGCTGGCA 903

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              |||||||||||||| ||||| || |||||||| ||||||||||||||||||||||||||
Db        904 GAGCTGGTGCGGTCTGTGCGCCTGGTGGAGCGGGCCCTGGGCTCCCCAACCAAGCAGCTG 963

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||
Db        964 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTCGGCAAGTCTGTGGTAGCCAAAGTGAAA 1023

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              || || | |||||||| || || |||||||||||||| |||||||||| |||||||||
Db       1024 ATCCCAGCAGGCACCACCCTGACCCTGGACATGCTCACTGTGAAGGTGGGGGAGCCCAAA 1083

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | |||||||||||||||||||| || |||| |||||| ||||| ||||||||| | |||
Db       1084 GGCTATCCTCCTGAAGACATCTTCAACCTAGCGGGCAAAAAGGTGCTGGTCACTATCGAA 1143

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              || |||||||| |||||||||||| || || | ||| |||| ||||||||| |||||
Db       1144 GAAGATGACACGGTCATGGAAGAATCCGTGGAAAGTCACAGCAAGAAAATCAAGGCTTAA 1203
```

```
RESULT 10
BC057977
LOCUS       BC057977                1916 bp    mRNA    linear   ROD 08-OCT-2003
DEFINITION  Mus musculus N-acetylneuraminic acid synthase (sialic acid
            synthase), mRNA (cDNA clone MGC:67938 IMAGE:5052339), complete cds.
ACCESSION   BC057977
VERSION     BC057977.1  GI:37046731
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1916)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1916)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 123 Row: f Column: 14
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 16716466.
FEATURES             Location/Qualifiers
     source          1. .1916
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="MGC:67938 IMAGE:5052339"
                     /tissue_type="Liver, normal. 5 month old male mouse."
                     /clone_lib="NCI_CGAP_Li9"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .1916
                     /gene="Nans"
                     /db_xref="LocusID:94181"
                     /db_xref="MGI:2149820"
     CDS             5. .1084
                     /codon_start=1
                     /product="N-acetylneuraminic acid synthase (sialic acid
                     synthase)"
                     /protein_id="AAH57977.1"
                     /db_xref="GI:37046732"
                     /db_xref="LocusID:94181"
                     /translation="MPLELELCPGRWVGGKHPCFIIAEIGQNHQGDIDVAKRMIRTAK
                     ECGADCAKFQKSELEFKFNRKALERPYTSKHSWGKTYGEHKRHLEFSHDQYKELQSYA
                     QEIGIFFTASGMDEMAVEFLHELNVPFFKVGSGDTNNFPYLEKTAKKGRPMVISSGMQ
                     SMDTMKQVYQIVKPLNPNFCFLQCTSAYPLQPEDANLRVISEYQKLFPDIPIGYSGHE
                     TGIAISVAAVALGAKVLERHITLDKTWKGSDHSASLEPGELAELVRSVRLVERALGSP
                     TKQLLPCEMACNEKLGKSVVAKVKIPAGTTLTLDMLTVKVGEPKGYPPEDIFNLAGKK
                     VLVTIEEDDTVMEESVESHSKKIKA"
```

```
     misc_feature    26. .1057
                     /note="SpsE; Region: Sialic acid synthase [Cell envelope
                     biogenesis, outer membrane]"
                     /db_xref="CDD:COG2089"
ORIGIN

  Query Match           81.8%;  Score 883.2;  DB 10;  Length 1916;
  Best Local Similarity   88.6%;  Pred. No. 1.2e-207;
  Matches  957;  Conservative    0;  Mismatches  123;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              |||||||||| |||||||||||||||||||||||||||||| ||  | ||||||||||||
Db          5 ATGCCGCTGGAACTGGAGCTGTGTCCCGGGCGCTGGGTGGGTGGAAAGCACCCGTGCTTC 64

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||| || |||||||||||||||||||| || ||| | || || ||||||||||||||||
Db         65 ATCATCGCGGAGATCGGCCAGAACCACCAAGGAGACATAGATGTGGCCAAGCGCATGATC 124

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||  |||||||||||||||||| || || || ||||| |||||||| ||| | || |||
Db        125 CGCACTGCCAAGGAGTGTGGGGCCGACTGCGCTAAGTTTCAGAAGAGCGAGTTGGAGTTC 184

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||| ||||| ||| ||||||| ||||| || |||||||||| ||||||||||||
Db        185 AAGTTTAACCGGAAGGCCCTGGAGAGACCATATACTTCGAAGCATTCATGGGGGAAGACG 244

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              || |||||||||| || |||||||| |||||||| |||||||||| ||||||||||||
Db        245 TATGGGGAGCACAAGCGGCATCTGGAATTCAGCCACGACCAGTACAAGGAGCTGCAGAGC 304

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              || ||  ||||| | || |||||||||||||||||||||||||||| ||||||||||||||
Db        305 TATGCGCAGGAGATCGGCATCTTCTTCACTGCCTCTGGCATGGACGAGATGGCAGTTGAG 364

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              || ||||| |||||||||||||| |||||||||||||||||||| ||||||| ||||||
Db        365 TTTCTGCACGAACTGAATGTTCCCTTTTTCAAAGTTGGATCTGGGGACACTAACAATTTT 424

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              || || |||||||||||||||| |||||||| || |||||||||||||| |||||||||
Db        425 CCCTACCTGGAAAAGACAGCCAAGAAAGGTCGTCCTATGGTGATCTCCAGCGGGATGCAG 484

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              |||||||||||||||||||||||||||||||||||||||||||| || ||||||||||||
Db        485 TCAATGGACACCATGAAGCAAGTCTATCAGATCGTGAAGCCGCTGAATCCCAACTTCTGC 544

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||| | || || |||||||| ||||| || ||||| ||||| | ||||||||| ||||||
Db        545 TTCCTCCAATGCACCAGCGCGTACCCACTACAGCCCGAGGATGCCAACCTGCGCGTCATC 604

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              |||||||| |||||||||||||| |||||||||| |||||||| ||||| || || |||
Db        605 TCGGAATACCAGAAGCTCTTTCCCGACATTCCCATCGGGTATTCCGGGCACGAGACGGGC 664

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              || || ||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        665 ATCGCCATATCTGTGGCCGCCGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACG 724

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              |||||||||||||||||||||||||||||||| |||||||||||||||| || |||||
Db        725 TTGGACAAGACCTGGAAGGGGAGTGACCACTCAGCCTCGCTGGAGCCTGGGGAGCTGGCA 784

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              |||||||||||||| ||||| || |||||||| ||||||||||||||||||||||||||
Db        785 GAGCTGGTGCGGTCTGTGCGCCTGGTGGAGCGGGCCCTGGGCTCCCCAACCAAGCAGCTG 844

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db        845 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTCGGCAAGTCTGTGGTAGCCAAAGTGAAA 904

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              || || | ||||||||  || ||  |||||||||||| |||||||||| |||||||||
Db        905 ATCCCAGCAGGCACCACCCTGACCCTGGACATGCTCACTGTGAAGGTGGGGGAGCCCAAA 964
```

```
Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | |||||||||||||||||||||||| || |||| |||||| ||||| ||||||||| | |||
Db        965 GGCTATCCTCCTGAAGACATCTTCAACCTAGCGGGCAAAAAGGTGCTGGTCACTATCGAA 1024

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              || ||||||||  |||||||||||  || || | |||  |||| |||||||||| |||||
Db       1025 GAAGATGACACGGTCATGGAAGAATCCGTGGAAAGTCACAGCAAGAAAATCAAGGCTTAA 1084


RESULT 11
AB041263
LOCUS       AB041263                1174 bp    mRNA    linear   ROD 15-JUL-2000
DEFINITION  Mus musculus mRNA for N-acetylneuraminic acid 9-phosphate
            synthetase, complete cds.
ACCESSION   AB041263
VERSION     AB041263.1  GI:8978233
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (sites)
  AUTHORS   Nakata,D., Close,B.E., Colley,K.J., Matsuda,T. and Kitajima,K.
  TITLE     Molecular cloning and expression of the mouse N-acetylneuraminic
            acid 9-phosphate synthase which does not have deaminoneuraminic
            acid (KDN) 9-phosphate synthase activity
  JOURNAL   Biochem. Biophys. Res. Commun. 273 (2), 642-648 (2000)
  MEDLINE   20334323
   PUBMED   10873658
REFERENCE   2  (bases 1 to 1174)
  AUTHORS   Kitajima,K. and Nakata,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-APR-2000) Ken Kitajima, Graduate school of
            bioagricultural sciences,Nagoya University, Department of applied
            Molecular biosciences; Chikusa-ku Hurou-chou, Nagoya, Aichi
            464-8601, Japan (E-mail:kitajima@agr.nagoya-u.ac.jp,
            Tel:+81-52-789-4130, Fax:+81-52-789-4128)
FEATURES             Location/Qualifiers
     source          1. .1174
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /cell_line="COMMA1d"
     CDS             71. .1150
                     /codon_start=1
                     /product="N-acetylneuraminic acid 9-phosphate synthetase"
                     /protein_id="BAA98131.1"
                     /db_xref="GI:8978234"
                     /translation="MPLELELCPGRWVGGKHPCFIIAEIGQNHQGDIDVAKRMIRTAK
                     ECGADCAKFQKSELEFKFNRKALERPYTSKHSWGKTYGEHKRHLEFSHDQYKELQSYA
                     QEIGIFFTASGMDEMAVEFLHELNVPFFKVGSGDTNNFPYLEKTAKKGRPMVISSGMQ
                     SMDTMKQVYQIVKPLNPNFCFLQCTSAYPLQPEDANLRVISEYQKLFPDIPIGYSGHE
                     TGIAISVAAVALGAKVLERHITLDKTWKGSDHSASLEPGELAELVRSVRLVERALGSP
                     TKQLLPCEMACNEKLGKSVVAKVKIPAGTTLTLDTLTVKVGEPKGYPPEDIFNLAGKK
                     VLVTIEEDDTVMEESVESHSKKIKA"
ORIGIN

  Query Match           81.6%;  Score 881.6;  DB 10;  Length 1174;
  Best Local Similarity   88.5%;  Pred. No. 2.9e-207;
  Matches  956;  Conservative    0;  Mismatches  124;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              |||||||||| |||||||||||||||||||||||||||||| ||  | ||||||||||||
Db         71 ATGCCGCTGGAACTGGAGCTGTGTCCCGGGCGCTGGGTGGGTGGAAAGCACCCGTGCTTC 130

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||| || |||||||||||||||||||| || ||| | || || ||||||||||||||||
Db        131 ATCATCGCGGAGATCGGCCAGAACCACCAAGGAGACATAGATGTGGCCAAGCGCATGATC 190

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||  |||||||||||||||||| || || || ||||| |||||||| ||| | || |||
Db        191 CGCACTGCCAAGGAGTGTGGGGCCGACTGCGCTAAGTTTCAGAAGAGCGAGTTGGAGTTC 250
```

```
Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
Db        251 AAGTTTAACCGGAAGGCCCTGGAGAGACCATATACTTCGAAGCATTCATGGGGGAAGACG 310

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
Db        311 TATGGGGAGCACAAGCGGCATCTGGAATTCAGCCACGACCAGTACAAGGAGCTGCAGAGC 370

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
Db        371 TATGCGCAGGAGATCGGCATCTTCTTCACTGCCTCTGGCATGGACGAGATGGCAGTTGAG 430

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
Db        431 TTTCTGCACGAACTGAATGTTCCCTTTTTCAAAGTTGGATCTGGGGACACTAACAATTTT 490

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
Db        491 CCCTACCTGGAAAAGACAGCCAAGAAAGGTCGTCCTATGGTGATCTCCAGCGGGATGCAG 550

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
Db        551 TCAATGGACACCATGAAGCAAGTCTATCAGATCGTGAAGCCGCTGAATCCCAACTTCTGC 610

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
Db        611 TTCCTCCAATGCACCAGCGCGTACCCACTACAGCCCGAGGATGCCAACCTGCGCGTCATC 670

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
Db        671 TCGGAATACCAGAAGCTCTTTCCCGACATTCCCATCGGGTATTCCGGGCACGAGACGGGC 730

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
Db        731 ATCGCCATATCTGTGGCCGCCGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACG 790

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
Db        791 TTGGACAAGACCTGGAAGGGGAGTGACCACTCAGCCTCGCTGGAGCCTGGGGAGCTGGCA 850

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
Db        851 GAGCTGGTGCGGTCTGTGCGCCTGGTGGAGCGGGCCCTGGGCTCCCCAACCAAGCAGCTG 910

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
Db        911 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTCGGCAAGTCTGTGGTAGCCAAAGTGAAA 970

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
Db        971 ATCCCAGCAGGCACCACCCTGACCCTGGACACGCTCACTGTGAAGGTGGGGGAGCCCAAA 1030

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
Db       1031 GGCTATCCTCCTGAAGACATCTTCAACCTAGCGGGCAAAAAGGTGCTGGTCACTATCGAA 1090

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
Db       1091 GAAGATGACACGGTCATGGAAGAATCCGTGGAAAGTCACAGCAAGAAAATCAAGGCTTAA 1150
```

RESULT 12
AK091184
LOCUS AK091184 3026 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ33865 fis, clone CTONG2006515, highly similar to N-acetylneuraminic acid phosphate synthase.
ACCESSION AK091184
VERSION AK091184.1 GI:21749493
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Nishi,T., Ota,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3026)
AUTHORS Isogai,T. and Yamamoto,J.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
FEATURES Location/Qualifiers
source 1. .3026
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CTONG2006515"
/tissue_type="tongue, tumor tissue"
/clone_lib="CTONG2"
/note="cloning vector: pME18SFL3"
ORIGIN

Query Match 67.8%; Score 732.4; DB 9; Length 3026;
Best Local Similarity 99.9%; Pred. No. 2.7e-170;
Matches 733; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        347 AGATGGCAGTTGAATTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAG 406
Db       2273 AGATGGCAGTTGAATTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAG 2332

Qy        407 ACACTAATAATTTTCCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCT 466
Db       2333 ACACTAATAATTTTCCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCT 2392

Qy        467 CCAGTGGGATGCAGTCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCA 526
Db       2393 CCAGTGGGATGCAGTCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCA 2452

Qy        527 ACCCCAACTTCTGCTTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCA 586
Db       2453 ACCCCAACTTCTGCTTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCA 2512

Qy        587 ACCTGCGGGTCATCTCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTG 646
Db       2513 ACCTGCGGGTCATCTCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTG 2572

Qy        647 GGCATGAAACAGGCATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGG 706
Db       2573 GGCATGAAACAGGCATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGG 2632

Qy        707 AACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGC 766
Db       2633 AACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGC 2692

Qy        767 CTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCC 826
Db       2693 CTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCC 2752

Qy        827 CAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGG 886
```

```
Db       2753 CAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGG 2812

Qy        887 TGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGG 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 TGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGG 2872

Qy        947 TGGGTGAGCCCAAAGCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCC 1006
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       2873 TGGGTGAGCCCAAAGGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCC 2932

Qy       1007 TGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAA 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 TGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAA 2992

Qy       1067 AAATCAAGTCTTAA 1080
              ||||||||||||||
Db       2993 AAATCAAGTCTTAA 3006
```

RESULT 13
AX868777

```
LOCUS       AX868777                 813 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 3682 from Patent EP1074617.
ACCESSION   AX868777
VERSION     AX868777.1  GI:40023640
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 3682 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .813
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
  Query Match             65.8%;  Score 710.2;  DB 6;  Length 813;
  Best Local Similarity   98.7%;  Pred. No. 7.8e-165;
  Matches  734;  Conservative    0;  Mismatches    8;  Indels    2;  Gaps    2;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         71 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGNCAACACCCGTGCTTC 130

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 190

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 250

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        251 AAGTTTAATCGGAAAGCCTTGGACAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 310

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 370

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        371 TACGCCGAGGAGGTTGGGATCTTCTTCACTGNCTCTGGCATGGATGAGATGGCAGTTGAA 430

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 490

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        491 CCTTATCTGGAAAAGACAGNCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 550

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 610

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 670

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        671 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 730

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGC-TCTGGGGGCCAAGGTGTTGGAACGTCACATAAC 719
              |||||||||||||||||||||||||| | ||||| |||||||| |||||||||||| ||
Db        731 ATAGCGATATCTGTGGCCGCAGTGGCTTTTGGGGACCAAGGTGNTGGAACGTCACAT-AC 789

Qy        720 TTTGGACAAGACCTGGAAGGGGAG 743
              |||||||||||||||||| |||||
Db        790 TTTGGACAAGACCTGGAANGGGAG 813
```

RESULT 14
BD148839

```
LOCUS       BD148839                 813 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD148839
VERSION     BD148839.1  GI:27854597
KEYWORDS    JP 2002191363-A/3682.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 813)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 3682 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/3682
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                     Location/Qualifiers
            FT   source          1. .813
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .813
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
  Query Match             65.8%;  Score 710.2;  DB 6;  Length 813;
  Best Local Similarity   98.7%;  Pred. No. 7.8e-165;
  Matches  734;  Conservative    0;  Mismatches    8;  Indels    2;  Gaps    2;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
```

```
Db         71 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGNCAACACCCGTGCTTC 130

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 190

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 250

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        251 AAGTTTAATCGGAAAGCCTTGGACAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 310

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 370

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        371 TACGCCGAGGAGGTTGGGATCTTCTTCACTGNCTCTGGCATGGATGAGATGGCAGTTGAA 430

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 490

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        491 CCTTATCTGGAAAAGACAGNCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 550

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 610

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 670

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        671 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 730

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGC-TCTGGGGGCCAAGGTGTTGGAACGTCACATAAC 719
              |||||||||||||||||||||||| | ||||| |||||||| |||||||||||||| ||
Db        731 ATAGCGATATCTGTGGCCGCAGTGGCTTTTGGGGACCAAGGTGNTGGAACGTCACAT-AC 789

Qy        720 TTTGGACAAGACCTGGAAGGGGAG 743
              |||||||||||||||||| |||||
Db        790 TTTGGACAAGACCTGGAANGGGAG 813


RESULT 15
AX873806/c
LOCUS       AX873806                514 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 8711 from Patent EP1074617.
ACCESSION   AX873806
VERSION     AX873806.1  GI:40028595
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 8711 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .514
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             44.8%;  Score 483.8;  DB 6;  Length 514;
  Best Local Similarity   98.2%;  Pred. No. 8.6e-109;
  Matches  485;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy        587 ACCTGCGGGTCATCTCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTG 646
              || || |||| | | |||| ||||||||||||||||||||||||||||||||||||||||
Db        514 ACNTGNGGGTNNTNTNGGAANATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTG 455

Qy        647 GGCATGAAACAGGCATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGG 706
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        454 GGCATGAAACAGGCATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGACCAAGGTGTTGG 395

Qy        707 AACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGC 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 AACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGC 335

Qy        767 CTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCC 826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 CTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCC 275

Qy        827 CAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGG 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 CAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGG 215

Qy        887 TGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGG 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 TGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGG 155

Qy        947 TGGGTGAGCCCAAAGCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCC 1006
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        154 TGGGTGAGCCCAAAGGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCC 95

Qy       1007 TGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAA 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 TGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAA 35

Qy       1067 AAATCAAGTCTTAA 1080
              ||||||||||||||
Db         34 AAATCAAGTCTTAA 21


Search completed: September 16, 2004, 04:04:22
Job time : 6858 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 15, 2004, 21:22:11 ; Search time 668 Seconds
(without alignments)
6868.344 Million cell updates/sec

Title: US-10-759-277-3
Perfect score: 1080
Sequence: 1 atgccgctggagctggagct.....gcaaaaaaatcaagtcttaa 1080

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1080	100.0	1080	3	AAA37763	Aaa37763 Human gly
	2	1080	100.0	1080	4	AAH28458	Aah28458 Nucleotid
	3	1080	100.0	1080	7	AAL53993	Aal53993 DNA encod
	4	1080	100.0	1146	3	AAA50569	Aaa50569 Human sia
	5	1078	99.8	1268	6	ABL90630	Abl90630 Human pol
	6	1076.8	99.7	1230	4	AAI59142	Aai59142 Human pol
	7	1076.8	99.7	1230	8	ADB49125	Adb49125 Novel hum
	8	1075.2	99.6	1170	4	AAH14557	Aah14557 Human cDN
	9	898.2	83.2	1243	4	AAI60928	Aai60928 Human pol
c	10	732	67.8	1238	4	AAS33141	Aas33141 DNA encod
	11	710.2	65.8	813	4	AAH06847	Aah06847 Human cDN
c	12	483.8	44.8	514	4	AAH11876	Aah11876 Human cDN
	13	426.4	39.5	1417	5	AAS91837	Aas91837 DNA encod
c	14	418	38.7	464	6	ABN95016	Abn95016 Gene #151
c	15	386.8	35.8	553	6	ABK94731	Abk94731 Neurodege
	16	371.8	34.4	486	8	ACH18008	Ach18008 Human adu
	17	336.8	31.2	442	8	ACH41408	Ach41408 Human foe
	18	327.8	30.4	417	8	ACH38131	Ach38131 Human end
	19	315.8	29.2	480	8	ACH32141	Ach32141 Human end
	20	280.6	26.0	404	5	AAH42798	Aah42798 Nucleotid
	21	260.6	24.1	1028	5	AAS81438	Aas81438 DNA encod
	22	256.2	23.7	410	4	AAH35557	Aah35557 Human col
	23	198.2	18.4	1017	4	ABL29095	Abl29095 Drosophil
	24	162.4	15.0	611	5	AAS91836	Aas91836 DNA encod
c	25	146	13.5	492	4	AAK88570	Aak88570 Human dig
	26	119.8	11.1	3123	4	ABL29094	Abl29094 Drosophil
	27	119.8	11.1	3375	4	ABL29196	Abl29196 Drosophil
	28	112.4	10.4	26281	3	AAZ60929	Aaz60929 Nucleotid
	29	101.6	9.4	17276	7	ACA64723	Aca64723 Streptoco
	30	100	9.3	1023	6	ABN67441	Abn67441 Streptoco
c	31	100	9.3	110000	6	ABN71527_11	Continuation (12 o
	32	90	8.3	3777	4	ABL29200	Abl29200 Drosophil
	33	70.4	6.5	110000	2	AAV21209_04	Continuation (5 of
	34	63	5.8	1059	3	AAA50570	Aaa50570 Escherich
	35	63	5.8	1059	4	AAH28459	Aah28459 Nucleotid
	36	63	5.8	1059	7	AAL53994	Aal53994 DNA encod
	37	62.8	5.8	999	4	AAF91405	Aaf91405 N. mening
	38	62.8	5.8	1000	6	ABK37785	Abk37785 DNA seque
	39	62.8	5.8	1050	6	ABA99765	Aba99765 N. mening
	40	62.8	5.8	65632	3	AAA81502	Aaa81502 N. mening
c	41	62.8	5.8	110000	3	AAA81490_00	Aaa81490 N. mening
c	42	62.8	5.8	349980	3	AAF21544	Aaf21544 Neisseria
	43	62	5.7	133	3	AAC98691	Aac98691 Human col
	44	60	5.6	60	6	ABN36327	Abn36327 Human spl
	45	59.6	5.5	984	6	ABV72354	Abv72354 Nucleotid

ALIGNMENTS

RESULT 1
AAA37763
ID AAA37763 standard; DNA; 1080 BP.
XX
AC AAA37763;
XX
DT 04-DEC-2000 (first entry)
XX
DE Human glycosylation enzyme clone HA5AA37 coding sequence.
XX
KW Human; glycosylation enzyme; glycolysis; myoglobinuria; tumour marker;
KW immunotherapy; cosmetic surgery; metabolism; immune system disorder;
KW haematopoietic cell deficiency; blood coagulation disorder; asthma;
KW afibrinogenaemia; blood platelet disorder; thrombocytopaenia; neoplasia;
KW autoimmune disorder; Addison's disease; multiple sclerosis; purpura;
KW allergic encephalomyelitis; allergic reaction; organ rejection;
KW graft-versus-host disease; inflammation; hyperproliferative disorder;
KW sarcoidosis; infection; gene therapy; CMP sialic acid synthetase; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .1080
FT /*tag= a
FT /product= "Sialic_acid_synthetase"
XX
PN WO200052136-A2.
XX
PD 08-SEP-2000.
XX
PF 01-MAR-2000; 2000WO-US005325.
XX
PR 02-MAR-1999; 99US-0122409P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Coleman TA;
XX
DR WPI; 2000-572179/53.
DR P-PSDB; AAY90352.
XX
PT New human glycosylation enzymes cytidine 5'-monophosphate sialic acid
PT synthetase, sialic acid synthetase and aldolase and nucleic acids
PT encoding the proteins for treating e.g., immune system disorders,
PT microbial diseases.
XX

```
PS   Claim 4; Page 110-111; 115pp; English.
XX
CC   This sequence encodes a human glycosylation enzyme clone of the
CC   invention, designated HA5AA37. The protein of this clone is a sialic acid
CC   synthetase. The sequences are useful as reagents for the differential
CC   identification of the tissues or cell types present in a biological
CC   sample, as immunological probes, for treating a disease or condition
CC   resulting from under expression of such polypeptide, for the detection
CC   and/or treatment of disorders involving aberrant glycolysis, e.g. cramps,
CC   myoglobinuria, and as tumour marker and/or immunotherapy targets. They
CC   may also be used to differentiate, proliferate and attract cells leading
CC   to the regeneration of tissues, to modulate mammalian characteristics
CC   (e.g. in cosmetic surgery) or mammalian metabolism affecting catabolism,
CC   anabolism processing, utilisation and energy storage, to change a
CC   mammal's mental state by influencing and as a food additive or
CC   preservative. The proteins can be used to assay protein levels in a
CC   sample, as a marker or detector of an immune system disorder, to inhibit
CC   cytokine activity, and as a vaccine. They may further be used to treat
CC   immune system or of haematopoietic cell deficiencies or disorders, blood
CC   coagulation disorders (e.g. afibrinogenaemia), blood platelet disorders
CC   (e.g. thrombocytopaenia), wounds resulting from trauma or surgery,
CC   autoimmune disorders (e.g. Addison's disease, multiple sclerosis,
CC   allergic encephalomyelitis), allergic reactions (e.g. asthma), organ
CC   rejection, graft-versus-host disease, inflammation, hyperproliferative
CC   disorders (e.g. neoplasia, purpura, sarcoidosis), diseases caused by
CC   viruses (e.g. hepatitis, meningitis, AIDS), bacteria and fungi (include
CC   e.g. tuberculosis, conjunctivitis, sepsis, typhoid, chlamydia,
CC   cellulitis), and diseases caused by parasites (e.g. amoebiasis,
CC   coccidiosis, leishmaniasis, scabies, malaria, toxoplasmosis)
XX
SQ   Sequence 1080 BP; 274 A; 261 C; 314 G; 231 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 1080;  DB 3;  Length 1080;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1080;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080

RESULT 2
AAH28458
ID   AAH28458 standard; DNA; 1080 BP.
XX
AC   AAH28458;
XX
DT   17-SEP-2001  (first entry)
XX
DE   Nucleotide sequence of a human sialic acid synthetase.
XX
KW   Carbohydrate processing; sialylation; sialylated glycoprotein; CMP-SA;
KW   cytidine monophosphate-sialic acid; N-acetylglucosamine-2 epimerase;
KW   sialic acid synthetase; aldolase; CMP-SA synthetase; CMP-SA transporter;
KW   vaccine; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1080
FT                   /*tag=  a
FT                   /product= "sialic acid synthetase"
XX
PN   WO200142492-A1.
XX
PD   14-JUN-2001.
XX
PF   07-DEC-2000; 2000WO-US033136.
XX
PR   09-DEC-1999;   99US-0169839P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (UTEM ) UNIV TEMPLE.
PA   (UYWY-) UNIV WYOMING.
XX
PI   Betenbaugh MJ,  Lawrence S,  Lee YC,  Coleman TA,  Palter K;
PI   Jarvis D;
```

```
XX
DR   WPI; 2001-441575/47.
DR   P-PSDB; AAB84683.
XX
PT   Cells producing cytidine monophosphate-sialic acid and sialylated
PT   glycoprotein above endogenous levels for production of vaccines and
PT   therapeutics.
XX
PS   Example 5; Fig 31; 182pp; English.
XX
CC   The specification describes a method for manipulating carbohydrate
CC   processing pathways in cells of interest. The methods are used to
CC   manipulate multiple pathways involved with the sialylation reaction by
CC   using recombinant DNA technology and substrate feeding approaches to
CC   enable the production of sialylated glycoproteins in the cells. The
CC   sialylation process involves the post-translational addition of the donor
CC   substrate cytidine monophosphate-sialic acid (CMP-SA) onto a specific
CC   acceptor carbohydrate. The cells express at least one enzyme, selected
CC   from N-acetylglucosamine-2 epimerase, sialic acid synthetase, aldolase,
CC   CMP-SA synthetase and CMP-SA transporter, above endogenous levels. The
CC   cells are useful for producing complex sialylated glycoproteins in cells
CC   of interest, especially insect cells. Glycoproteins containing sialylated
CC   oligosaccharides are useful as vaccines, therapeutics and diagnostic
CC   tools. Cells producing complex sialylated glycoproteins are useful for
CC   enhancing the value of heterologous expression systems and increasing the
CC   application of heterologous cell expression products as vaccines,
CC   therapeutics and diagnostic tools as well as increasing the variety of
CC   heterologous proteins that can be produced and lowering biotechnology
CC   production costs. The present sequence encodes a human sialic acid
CC   synthetase, which is used in the method of the invention
XX
SQ   Sequence 1080 BP; 274 A; 261 C; 314 G; 231 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 1080;  DB 4;  Length 1080;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1080;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080

RESULT 3
AAL53993
ID   AAL53993 standard; DNA; 1080 BP.
XX
AC   AAL53993;
XX
DT   06-MAR-2003  (first entry)
XX
DE   DNA encoding a human SA synthetase protein.
XX
KW   Vaccine; glycoprotein; insect cell; enzyme; N-acetylglucosamine-2;
KW   GlcNAc-2; epimerase; UDP-GlcNAc; mannose; (Man)NAc; sialic acid;
KW   synthetase; aldolase; cytidine monophosphate-sialic acid; CMP-SA;
KW   transporter; sialylated glycoprotein; human; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1080
FT                   /*tag=  a
FT                   /product= "Human SA synthetase protein"
XX
PN   US2002142386-A1.
XX
PD   03-OCT-2002.
XX
PF   16-AUG-2001; 2001US-00930440.
XX
PR   02-MAR-1999;   99US-0122582P.
PR   08-DEC-1999;   99US-0169624P.
PR   25-AUG-2000; 2000US-0227579P.
XX
PA   (BETE/) BETENBAUGH M J.
PA   (LAWR/) LAWRENCE S.
PA   (LEEY/) LEE Y C.
PA   (COLE/) COLEMAN T A.
```

```
XX
PI   Betenbaugh MJ,  Lawrence S,  Lee YC,  Coleman TA;
XX
DR   WPI; 2003-102519/09.
DR   P-PSDB; AAO26546.
XX
PT   Manipulating glycoprotein production in insect cell, involves enhancing
PT   expression of enzymes involved in carbohydrate processing pathway such as
PT   N-acetylglucosamine-2 epimerase or sialic acid synthetase.
XX
PS   Disclosure; Fig 31; 88pp; English.
XX
CC   The invention relates to a novel method for manipulating glycoprotein
CC   production in an insect cell comprising enhancing expression of an
CC   enzyme, such as N-acetylglucosamine-2 (GlcNAc-2) epimerase, one
CC   catalysing conversion of UDP-GlcNAc to mannose (Man)NAc, sialic acid
CC   synthetase, aldolase, cytidine monophosphate-sialic acid (CMP-SA)
CC   synthetase or CMP-SA transporter, where the expression of each enzyme is
CC   enhanced to above endogenous levels. The novel method is useful for
CC   manipulating glycoprotein production in an insect cell. Further methods
CC   of the invention are useful for producing sialylated glycoprotein. The
CC   sialylated glycoprotein produced by the above mentioned methods are
CC   useful as pharmaceutical compositions, vaccines, diagnostics and
CC   therapeutics. This polynucleotide sequence represents the DNA encoding a
CC   human SA synthetase protein of the invention
XX
SQ   Sequence 1080 BP; 274 A; 261 C; 314 G; 231 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 1080;  DB 7;  Length 1080;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1080;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080

RESULT 4
AAA50569
ID   AAA50569 standard; cDNA; 1146 BP.
XX
AC   AAA50569;
XX
DT   19-DEC-2000  (first entry)
XX
DE   Human sialic acid synthetase cDNA.
XX
KW   Sialic acid synthetase; human; sas gene; sialylation; glycoprotein;
KW   plasminogen; transferrin; thyrotropin; Na+,K+-ATPase; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             46. .1125
FT                   /*tag=  a
XX
PN   WO200052135-A2.
XX
PD   08-SEP-2000.
XX
PF   01-MAR-2000; 2000WO-US005313.
XX
PR   02-MAR-1999;   99US-0122582P.
PR   08-DEC-1999;   99US-0169624P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (UYWY-) UNIV WYOMING.
XX
PI   Betenbaugh MJ,  Lawrence S,  Lee YC,  Jarvis D,  Coleman TA;
XX
DR   WPI; 2000-572178/53.
DR   P-PSDB; AAY96101.
XX
PT   Recombinant production of sialylated glycoproteins using cells in which
PT   the expression of enzymes, e.g. sialic acid synthetase, involved in the
PT   sialylation reaction has been altered.
XX
```

PS Disclosure; Page 103-105; 144pp; English.
XX
CC The present sequence is that of human cDNA encoding sialic acid
CC synthetase (see AAY96101), an enzyme that condenses ManNAc-6-P or Man-6-P
CC with PEP to form Neu5Ac and KDN phosphates, respectively. This novel
CC gene, termed SAS, was identified on the basis of homology with the
CC Escherichia coli neuB gene using a bioinformatics based approach.
CC Northern blots indicated ubiquitous transcription of the human synthetase
CC gene in a selection of tissues. The invention provides methods and
CC recombinantly engineered cells for producing glycoproteins having
CC sialylated oligosaccharides. The methods involve altering the expression
CC of enzymes involved in carbohydrate processing. A claimed cell producing
CC sialylated glycoprotein at above endogenous levels expresses at least 1
CC (preferably human) enzyme selected from GlcNAc-2 epimerase, an enzyme
CC catalyzing the conversion of UDP-GlcNAc to ManNAc, sialic acid
CC synthetase, aldolase, CMP-sialic acid synthetase and CMP-sialic acid
CC transporter at above endogenous levels. Endogenous N-
CC acetylglucosaminidase activity may be suppressed. A claimed method for
CC manipulating glycoprotein in an insect cell comprises enhancing the
CC expression of 1 of the above enzymes, and a claimed method for producing
CC sialylated glycoproteins involves expressing a heterologous protein
CC (especially plasminogen, transferrin, Na+,K+-ATPase or thyrotropin) in
CC the insect cell. Yeast, insect, fungal, plant and bacterial host cells
CC can be engineered to produce new forms of sialylated glycoproteins,
CC higher concentrations of sialylated glycoproteins and/or elevated
CC concentrations of donor substrates (e.g. nucleotide sugars) required for
CC sialylation
XX
SQ Sequence 1146 BP; 290 A; 280 C; 332 G; 244 T; 0 U; 0 Other;

Query Match 100.0%; Score 1080; DB 3; Length 1146;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         46 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 105

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 165

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 225

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 285

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 345

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 405

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 465

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 525

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 585

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 645

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 705

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 765

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        766 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 825

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 885

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        886 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 945

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        946 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 1005

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1065

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1066 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1125
```

RESULT 5
ABL90630
ID ABL90630 standard; cDNA; 1268 BP.
XX
AC ABL90630;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1192.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US016450.
XX
PR 19-MAY-2000; 2000US-0205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.
DR P-PSDB; ABB90221.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX
PS Claim 4; SEQ ID NO 1192; 2081pp + Sequence Listing; English.

```
XX
CC   The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC   (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. The genes are
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC   the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC   ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC   breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC   disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC   cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC   ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC   infectious diseases such as viral, bacterial, fungal and parasitic
CC   infections. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1268 BP; 328 A; 304 C; 364 G; 258 T; 0 U; 14 Other;

  Query Match             99.8%;  Score 1078;  DB 6;  Length 1268;
  Best Local Similarity   99.8%;  Pred. No. 2.1e-314;
  Matches 1078;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 180

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 240

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        241 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCYAAGTTCCAGAAGAGTGAGCTAGAATTC 300

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 360

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 420

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 480

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 540

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 600

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 660

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 720

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 780

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 840

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 900

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 960

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 1020

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 1080

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1140

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1200


RESULT 6
AAI59142
ID   AAI59142 standard; cDNA; 1230 BP.
XX
AC   AAI59142;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 1345.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US034263.
XX
PR   23-DEC-1999;   99US-00471275.
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
PR   20-JUN-2000; 2000US-00598042.
PR   19-JUL-2000; 2000US-00620312.
PR   03-AUG-2000; 2000US-00653450.
PR   14-SEP-2000; 2000US-00662191.
PR   19-OCT-2000; 2000US-00693036.
PR   29-NOV-2000; 2000US-00727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J,  Zhao QA;
PI   Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM39986.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders such
PT   as central nervous system injuries.
XX
PS   Claim 1; SEQ ID NO 1345; 10078pp; English.
```

```
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC   encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders. Note: The sequence data for this patent did not form
CC   part of the printed specification
XX
SQ   Sequence 1230 BP; 310 A; 309 C; 358 G; 253 T; 0 U; 0 Other;

  Query Match           99.7%;  Score 1076.8;  DB 4;  Length 1230;
  Best Local Similarity 99.8%;  Pred. No. 6.2e-314;
  Matches 1078;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 178

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 238

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 298

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        299 AAGTTTAATCGGAAAGCCTTGGACAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 358

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 418

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 478

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 538

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 598

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 658

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 718

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 778

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 838

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 898

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 958

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 1018

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 1078

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 GGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1138

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1198


RESULT 7
ADB49125
ID   ADB49125 standard; cDNA; 1230 BP.
XX
AC   ADB49125;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Novel human cDNA SEQ ID NO 1035.
XX
KW   ss; cancer; neurodegenerative disease; human.
XX
OS   Homo sapiens.
XX
PN   US2003104529-A1.
XX
PD   05-JUN-2003.
XX
PF   04-JAN-2002; 2002US-00037270.
XX
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
PR   19-JUL-2000; 2000US-00620312.
XX
PA   (ZHOU/) ZHOU P.
PA   (TANG/) TANG Y T.
PA   (LIUC/) LIU C.
PA   (ASUN/) ASUNDI V.
PA   (DRMA/) DRMANAC R T.
XX
PI   Zhou P,  Tang YT,  Liu C,  Asundi V,  Drmanac RT;
XX
DR   WPI; 2003-678194/64.
XX
PT   New polynucleotide, useful for treating diseases e.g., cancer or
PT   neurodegenerative diseases.
XX
PS   Claim 1; SEQ ID NO 1035; 99pp; English.
XX
CC   The invention relates to a polynucleotide comprising a sequence given in
CC   the specification, or its mature protein-coding portion, or its
CC   complement. The polynucleotide is useful for treating diseases e.g.,
CC   cancer or neurodegenerative diseases and many others listed in the
CC   specification. The present sequence represents a novel human cDNA. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=20030104529.
XX
SQ   Sequence 1230 BP; 310 A; 309 C; 358 G; 253 T; 0 U; 0 Other;
```

```
Query Match             99.7%;  Score 1076.8;  DB 8;  Length 1230;
Best Local Similarity   99.8%;  Pred. No. 6.2e-314;
Matches 1078;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 178

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 238

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 298

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        299 AAGTTTAATCGGAAAGCCTTGGACAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 358

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 418

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 478

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 538

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 598

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 658

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 718

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 778

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 838

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 898

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 958

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 1018

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 1078

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 GGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1138

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1198
```

```
RESULT 8
AAH14557
ID   AAH14557 standard; cDNA; 1170 BP.
XX
AC   AAH14557;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:12129.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   27-AUG-1999;   99JP-00300253.
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 8; SEQ ID NO 12129; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesising 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 1170 BP; 294 A; 285 C; 347 G; 244 T; 0 U; 0 Other;

  Query Match             99.6%;  Score 1075.2;  DB 4;  Length 1170;
  Best Local Similarity   99.7%;  Pred. No. 1.6e-313;
  Matches 1077;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 130

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 190

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 250

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 AAGTTTAATCGGAAAGCCTTGGACAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 310

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 370

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 430

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 490

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        491 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 550

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 610

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 670

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        671 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 730

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        731 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGACCAAGGTGTTGGAACGTCACATAACT 790

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        791 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 850

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        851 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 910

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        911 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 970

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        971 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 1030

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1031 GGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1090

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1091 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1150
```

RESULT 9
AAI60928
ID AAI60928 standard; cDNA; 1243 BP.
XX
AC AAI60928;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4917.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM41772.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 4917; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 1243 BP; 298 A; 291 C; 385 G; 269 T; 0 U; 0 Other;

Query Match 83.2%; Score 898.2; DB 4; Length 1243;
Best Local Similarity 88.6%; Pred. No. 4.5e-260;
Matches 1041; Conservative 0; Mismatches 3; Indels 131; Gaps 1;

```
Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        69 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 128

Qy        61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       129 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGATGTAGCCAAGCGCATGATC 188

Qy       121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       189 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 248

Qy       181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       249 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 308

Qy       241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       309 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 368

Qy       301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       369 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 428

Qy       361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       429 TTTCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 488

Qy       421 CCTTATCTGGAAAAGACAGCCAAAAA---------------------------------- 446
             |||||||||||||||||||||||||
Db       489 CCTTATCTGGAAAAGACAGCCAAAAAAGGGGCTGGCACAGTGTTTTAAGAGATGTCTGTG 548

Qy       447 ---------------------------------------------------------- 446

Db       549 GAGTTCAGTTGAATGATGAGACCAGCAGTTGGGATGTGTTGGGGAGAGTCAGAACCTCTA 608

Qy       447 -----------------------------------AGGTCGCCCAATGGTGATCTCCA 469
                                                 ||||||||||||||||||||||||
Db       609 AAGAGAAAGTGCTGATGGTGTTGGTGCTGGATTACTCAGGTCGCCCAATGGTGATCTCCA 668

Qy       470 GTGGGATGCAGTCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACC 529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       669 GTGGGATGCAGTCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACC 728

Qy       530 CCAACTTCTGCTTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACC 589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       729 CCAACTTCTGCTTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACC 788

Qy       590 TGCGGGTCATCTCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGC 649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       789 TGCGGGTCATCTCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGC 848

Qy       650 ATGAAACAGGCATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAAC 709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       849 ATGAAACAGGCATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAAC 908

Qy       710 GTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTG 769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       909 GTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTG 968

Qy       770 GAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAA 829
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       969 GAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAA 1028

Qy       830 CCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGG 889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1029 CCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGG 1088

Qy       890 CCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGG 949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1089 CCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGG 1148

Qy       950 GTGAGCCCAAAGCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGG 1009
             |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db      1149 GTGAGCCCAAAGGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGG 1208

Qy      1010 TCACTGTTGAAGAGGATGACACCATCATGGAAGAA 1044
             |||||||||||||||||||||||||||||||||||
Db      1209 TCACTGTTGAAGAGGATGACACCATCATGGAAGAA 1243
```

RESULT 10
AAS33141/c
ID AAS33141 standard; cDNA; 1238 BP.
XX
AC AAS33141;
XX
DT 04-DEC-2001 (first entry)
XX
DE DNA encoding human secreted protein, Seq ID No 100.
XX
KW Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
KW Gaucher's disease; neurological disease; cerebrovascular disorder;
KW thrombosis; wound healing; ss.
XX
OS Homo sapiens.
XX
PN WO200155326-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001347.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-451931/48.
DR P-PSDB; AAU20432.
XX
PT New nucleic acids and polypeptides, useful for diagnosing, preventing or
PT treating medical conditions.
XX
PS Claim 1; SEQ ID NO 100; 753pp; English.
XX
CC The invention relates to novel isolated nucleic acid molecules (I)
CC encoding human secreted proteins (II). (I) and (II) are used to prevent,
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
CC the prevention, treatment and diagnosis of diseases associated with
CC inappropriate expression of secreted proteins. (I) and complementary
CC sequences may also be used as DNA probes in diagnostic assays (e.g.
CC polymerase chain reactions (PCR)) to detect and quantitate the presence
CC of similar nucleic acid sequences in samples, and so which patients may
CC be in need of restorative therapy. (II) may also be used as antigens in
CC the production of antibodies and in assays to identify modulators
CC (agonists and antagonists) of the expression and activity of the secreted
CC proteins. The anti-(II) antibodies and antagonists may also be used to
CC down regulate expression and activity of (II). The anti-(II) antibodies
CC may also be used as diagnostic agents for detecting the presence of (II)
CC in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The
CC disorders include for example: immune/autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and
CC thrombosis), infections caused by bacteria, viruses and fungi and ocular
CC disorders (e.g. corneal infections). (I) and (II), agonists, antagonists
CC and antibodies can also be used to promote wound healing, maintain organs
CC before transplantation, and support cell culture of primary tissues.

Query Match 67.8%; Score 732; DB 4; Length 1238;
Best Local Similarity 99.7%; Pred. No. 6.5e-210;
Matches 732; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 347 AGATGGCAGTTGAATTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAG 406
||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db 799 AGATGGCAGTTGAATYCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAG 740

Qy 407 ACACTAATAATTTTCCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCT 466
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db        739 ACACTAATAATTTTCCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCT 680

Qy        467 CCAGTGGGATGCAGTCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCA 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 CCAGTGGGATGCAGTCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCA 620

Qy        527 ACCCCAACTTCTGCTTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCA 586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 ACCCCAACTTCTGCTTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCA 560

Qy        587 ACCTGCGGGTCATCTCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTG 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 ACCTGCGGGTCATCTCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTG 500

Qy        647 GGCATGAAACAGGCATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGG 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 GGCATGAAACAGGCATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGG 440

Qy        707 AACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGC 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 AACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGC 380

Qy        767 CTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCC 826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 CTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCC 320

Qy        827 CAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGG 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 CAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGG 260

Qy        887 TGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGG 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 TGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGG 200

Qy        947 TGGGTGAGCCCAAAGCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCC 1006
              ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        199 TGGGTGAGCCCAAAGGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCC 140

Qy       1007 TGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAA 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 TGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAA 80

Qy       1067 AAATCAAGTCTTAA 1080
              ||||||||||||||
Db          79 AAATCAAGTCTTAA 66
```

```
RESULT 11
AAH06847
ID   AAH06847 standard; cDNA; 813 BP.
XX
AC   AAH06847;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA clone (5'-primer) SEQ ID NO:3682.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   27-AUG-1999;   99JP-00300253.
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI   Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 1; SEQ ID NO 3682; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesising 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 813 BP; 198 A; 199 C; 243 G; 168 T; 0 U; 5 Other;

  Query Match           65.8%;  Score 710.2;  DB 4;  Length 813;
  Best Local Similarity 98.7%;  Pred. No. 2e-203;
  Matches  734;  Conservative    0;  Mismatches    8;  Indels    2;  Gaps    2;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         71 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGNCAACACCCGTGCTTC 130

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 190

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 250

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        251 AAGTTTAATCGGAAAGCCTTGGACAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 310

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 370

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        371 TACGCCGAGGAGGTTGGGATCTTCTTCACTGNCTCTGGCATGGATGAGATGGCAGTTGAA 430

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 490

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
```

```
Db        491 CCTTATCTGGAAAAGACAGNCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 550

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 610

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 670

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        671 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 730

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGC-TCTGGGGGCCAAGGTGTTGGAACGTCACATAAC 719
              |||||||||||||||||||||||||| | ||||| |||||||| |||||||||||||| ||
Db        731 ATAGCGATATCTGTGGCCGCAGTGGCTTTTGGGGACCAAGGTGNTGGAACGTCACAT-AC 789

Qy        720 TTTGGACAAGACCTGGAAGGGGAG 743
              |||||||||||||||||| |||||
Db        790 TTTGGACAAGACCTGGAANGGGAG 813
```

RESULT 12
AAH11876/c
ID AAH11876 standard; cDNA; 514 BP.
XX
AC AAH11876;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (3'-primer) SEQ ID NO:8711.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 3; SEQ ID NO 8711; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 514 BP; 112 A; 149 C; 113 G; 133 T; 0 U; 7 Other;

Query Match 44.8%; Score 483.8; DB 4; Length 514;
Best Local Similarity 98.2%; Pred. No. 3.3e-135;
Matches 485; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Qy        587 ACCTGCGGGTCATCTCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTG 646
              || || |||| | | |||| ||||||||||||||||||||||||||||||||||||||||||
Db        514 ACNTGNGGGTNNTNTNGGAANATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTG 455

Qy        647 GGCATGAAACAGGCATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGG 706
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        454 GGCATGAAACAGGCATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGACCAAGGTGTTGG 395

Qy        707 AACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGC 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 AACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGC 335

Qy        767 CTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCC 826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 CTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCC 275

Qy        827 CAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGG 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 CAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGG 215

Qy        887 TGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGG 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 TGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGG 155

Qy        947 TGGGTGAGCCCAAAGCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCC 1006
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        154 TGGGTGAGCCCAAAGGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCC 95

Qy       1007 TGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAA 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 TGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAA 35

Qy       1067 AAATCAAGTCTTAA 1080
              ||||||||||||||
Db         34 AAATCAAGTCTTAA 21
```

RESULT 13
AAS91837
ID AAS91837 standard; cDNA; 1417 BP.
XX
AC AAS91837;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #27641.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX

```
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG27650.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 27641; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1417 BP; 353 A; 336 C; 409 G; 319 T; 0 U; 0 Other;

  Query Match             39.5%;  Score 426.4;  DB 5;  Length 1417;
  Best Local Similarity   89.4%;  Pred. No. 1.2e-117;
  Matches  588;  Conservative    0;  Mismatches   46;  Indels   24;  Gaps   11;

Qy        447 AGGTCGCCCAATGGTGATCTCCAGTGGGATGCAGTCAATGGACACCATGAAGCAAGTTTA 506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 AGGTCGCCCAATGGTGATCTCCAGTGGGATGCAGTCAATGGACACCATGAAGCAAGTTAT 486

Qy        507 TCAGATCGTGAAGCCCCTCAACCCCAACTTCTGCTTC-TTGCAGTGTACCAGCGCATACC 565
              ||| |||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        487 TCAAATCGTGAAGCCCCTCAACCCCAACTTCTGCTTCTTTGCAGTGTACCAGCGCATACC 546

Qy        566 CGCTCCAGCCTGAGGACGTCAACC--TGCGGGTCATCTCGGAATATCAGAAGCTCTTTCC 623
              ||||||||||||||||||||||    ||||||||||||||||||||||||||||||||||
Db        547 CGCTCCAGCCTGAGGACGTCAACCCTGCCGGGTCATCTCGGAATATCAGAAGCTCTTTCC 606

Qy        624 TGACATTCCC---ATAGGGTATTCTGGGCATGAAACAGGCATAGCGATATCTGT---GGC 677
              ||||  | ||      || ||||||||||||||||||||     || | |   | |
Db        607 TGACCATTCCCATTAGGGTTATTCTGGGCATGAAACAGGGCATAGCCGATATCTTTGGCC 666

Qy        678 CGCAGTGGCTCTGGGGG-CCAAGGTGTTGGAA-----CGTCACATAACTTTGGACAAGAC 731
              | |||||||||||||| |||||||||||| |        || | |  || |||||||||
Db        667 CCCAGTGGCTCTGGGGGCCCAAGGTGTTGGGAACGTCTCACATAAACTTTGGGACAAGAC 726

Qy        732 CTGGAAGGGGAGTG----ACCACTCGGCCTCGCT--GGAGCCTGGAGAACTGGCCG-AGC 784
              | || | ||| | |     ||||||||||||||  |   | ||||||||||||| ||| 
Db        727 CGGGGAAGGGGGAGTTGACCCACTCGGCCTCGCTTGGAGCCTTGGAGAACTGGCCGAAGC 786

Qy        785 TGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACC-AAGCAGCTGCTG 843
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        787 TGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCTAAGCAGCTGCTG 846

Qy        844 CCCTGTGAGATGGCCTGCAA-TGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAAAT 902
              |||||||||||||||||||| |||||| ||||||||||||||||||||||||||||||
Db        847 CCCTGTGAGATGGCCTGCAATTGAGAACCTGGGCAAGTCTGTGGTGGCCAAAGTGAAAAT 906

Qy        903 TCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAAGC 962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 TCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAAGG 966

Qy        963 CTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAAGA 1022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 CTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAAGA 1026

Qy       1023 GGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 GGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1084


RESULT 14
ABN95016/c
ID   ABN95016 standard; DNA; 464 BP.
XX
AC   ABN95016;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Gene #1514 used to diagnose liver cancer.
XX
KW   Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW   metastatic liver tumour; cytostatic; expression profile; disease state;
KW   disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS   Homo sapiens.
XX
PN   WO200229103-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US030589.
XX
PR   02-OCT-2000; 2000US-0237054P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Horne D,  Alvares C,  Peres-Da-Silva S,  Vockley JG;
XX
DR   WPI; 2002-426119/45.
XX
PT   Diagnosing and detecting the progression of liver cancer, hepatocellular
PT   carcinoma or metastatic liver tumor in a patient, involves detecting the
PT   level of expression of two or more genes in a liver tissue sample.
XX
PS   Claim 1; SEQ ID NO 1514; 298pp; English.
XX
CC   The invention relates to a novel method for diagnosing and detecting the
CC   progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC   tumour in a patient, and differentiating metastatic liver cancer from
CC   hepatocellular carcinoma in a patient, involving detecting the level of
CC   expression of two or more genes represented in ABN93503-ABN97455 in a
CC   tissue sample. The method of the invention has hepatotropic, and
CC   cytostatic activity. The method is useful for diagnosing and detecting
CC   the progression of liver cancer, hepatocellular carcinoma and metastatic
CC   liver carcinoma in a patient. The method is useful for identifying
CC   expression profiles which serve as useful diagnostic markers as well as
CC   markers that can be used to monitor disease states, disease progression,
CC   drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
```

```
SQ   Sequence 464 BP; 101 A; 133 C; 103 G; 127 T; 0 U; 0 Other;

  Query Match             38.7%;  Score 418;  DB 6;  Length 464;
  Best Local Similarity   98.6%;  Pred. No. 2.2e-115;
  Matches  432;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy        643 TCTGGGCATGAAACAGGCATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTG 702
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||| ||
Db        464 TCTGGGCATGAAACAGGCATAGCGATATCTGTGGCCGCAGTGGCTCTTGGGGCCAAGTTG 405

Qy        703 TTGGAACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTG 762
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        404 TTGGAACGTCACATAACTTTGGACAAGACCTGGAAGTGGAGTGACCACTCGGCCTCGCTG 345

Qy        763 GAGCCTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGC 822
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        344 GAGCCTGGAGAACTGGCCGAGCTGGTGCGGTCATTGCGTCTTGTGGAGCGTGCCCTGGGC 285

Qy        823 TCCCCAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 TCCCCAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCT 225

Qy        883 GTGGTGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTG 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GTGGTGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTG 165

Qy        943 AAGGTGGGTGAGCCCAAAGCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAG 1002
              ||||||||||| ||||||| ||||||||||||||||||||||||||||||||||||||||
Db        164 AAGGTGGGTGA-CCCAAAGGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAG 106

Qy       1003 GTCCTGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        105 GTCCTGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGC 46

Qy       1063 AAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||
Db         45 AAAAAAATCAAGTCTTAA 28
```

```
RESULT 15
ABK94731/c
ID   ABK94731 standard; DNA; 553 BP.
XX
AC   ABK94731;
XX
DT   28-AUG-2002  (first entry)
XX
DE   Neurodegenerative disease associated polynucleotide #40.
XX
KW   Mitochondrial function; Forkhead Related Activator-2; FREAC-2; apoptosis;
KW   neurodegenerative disease; proliferative disorder; cancer;
KW   neurological disease; Alzheimer's disease; AD; multiple sclerosis; MS;
KW   Parkinson's disease; PD; vascular dementia; diffuse Lewy body disease;
KW   DLB; amylotrophic lateral sclerosis; ALS;
KW   myoclonic epilepsy lactic acidosis and stroke; MELAS; MERRF;
KW   myoclonic epilepsy ragged red fibre syndrome; autoimmune disease;
KW   gene therapy; antisense therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200240996-A2.
XX
PD   23-MAY-2002.
XX
PF   08-NOV-2001; 2001WO-US046699.
XX
PR   08-NOV-2000; 2000US-00709754.
XX
PA   (MITO-) MITOKOR.
XX
PI   Hwang J,  Becker DK,  Davis RE;
XX
DR   WPI; 2002-508336/54.
XX
PT   Identifying agents that alter mitochondrial function, in particular
PT   apoptosis, based on observation that expression of transcription factor,
PT   Forkhead Related Activator-2 in a cell results in apoptosis.
XX
PS   Disclosure; Page 150-151; 169pp; English.
XX
CC   The invention describes a method of identifying an agent that alters a
CC   mitochondrial function, comprising inducing Forkhead Related Activator-2
CC   (FREAC-2) expression in a host cell in the presence and absence of a
CC   candidate agent, and screening for an agent which alters the level of
CC   apoptosis resulting from such induced FREAC-2 expression. The method is
CC   useful for identifying an agent that alters a mitochondrial function,
CC   especially apoptosis. The invention also describes a method for
CC   identifying a molecular target for therapeutic intervention for treating
CC   a neurodegenerative disease or proliferative disorder. The methods of the
CC   invention are useful for treating a disease associated with inappropriate
CC   cell proliferation or cell survival, in particular cancer, or a
CC   neurological disease or disorder including Alzheimer's disease (AD),
CC   multiple sclerosis (MS), Parkinson's disease (PD), diffuse Lewy body
CC   disease (DLB), vascular dementia, amylotrophic lateral sclerosis (ALS),
CC   myoclonic epilepsy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC   ragged red fibre syndrome (MERRF) and other neurodegenerative diseases,
CC   and also autoimmune diseases. This sequence represents a polynucleotide
CC   associated with the neurodegenerative disorders described in the
CC   invention
XX
SQ   Sequence 553 BP; 121 A; 147 C; 137 G; 143 T; 0 U; 5 Other;

  Query Match             35.8%;  Score 386.8;  DB 6;  Length 553;
  Best Local Similarity   91.0%;  Pred. No. 6.4e-106;
  Matches  487;  Conservative    0;  Mismatches   37;  Indels   11;  Gaps    7;

Qy        555 CAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATCTCGGAATATCAGAA 614
              ||||      || | |||||||||||||||||||| ||||||||||||||||||||||||
Db        553 CAGCCATACCCGGTTCCAGCCTGAGGACGTCAACC-GCGGGTCATCTCGGAATATCAGAA 495

Qy        615 GCTCTTTCCTGACATTCCC---ATAGGGTATTCTGGCATGAAACAGGCATAGCGATATC 671
              |||||||||||| |  |     || |||||||| ||| ||||   ||         ||
Db        494 GCTCTTTCCTGACCATTNCCATTAGGGTATTCTGGCCATAAAACCAGCCNTAGCCGATA 435

Qy        672 TGT---GGCCGCAGTGGCTCTGGGGG-CCAAGGTGTTGGAACGTCACAT-AACTTTGGAC 726
              | |    | || |||||||||||||| ||||| |||||||||||||||| ||||||||||
Db        434 TCTTTGGCCCCCAGTGGCTCTGGGGGCCCAAGTTGTTGGAACGTCACATAAACTTTGGAC 375

Qy        727 AAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCCG-AGCT 785
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||| ||||
Db        374 AAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGANCCTGGAGAACTGGCCGAAGCT 315

Qy        786 GGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTGCTGCC 845
              ||| |||||||||||||||||||||| | ||||||||||||||||||||||||||||||
Db        314 GGTNCGGTCAGTGCGTCTTGTGGAGCNT-CCCTGGGCTCCCCAACCAAGCAGCTGCTGCC 256

Qy        846 CTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAAATTCC 905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 CTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAAATTCC 196

Qy        906 GGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAAGCCTA 965
              ||||||||||||||||||||||||||||||||||||||||||||||| |||||| |||
Db        195 GGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAACCCAAAGGCTA 136

Qy        966 TCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAAGAGGA 1025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAAGAGGA 76

Qy       1026 TGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 TGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 21
```

Search completed: September 16, 2004, 02:09:58
Job time : 674 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 16, 2004, 01:28:31 ; Search time 107 Seconds
(without alignments)
5601.373 Million cell updates/sec

Title: US-10-759-277-3
Perfect score: 1080
Sequence: 1 atgccgctggagctggagct.....gcaaaaaaatcaagtcttaa 1080

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1080	100.0	1080	4	US-09-516-143A-3	Sequence 3, Appli
	2	1076.8	99.7	1230	4	US-09-620-312D-1035	Sequence 1035, Ap
	3	73.6	6.8	1664976	4	US-08-916-421B-1	Sequence 1, Appli
	4	52.6	4.9	1234	4	US-08-976-259-81	Sequence 81, Appl
	5	44.4	4.1	11474	4	US-09-495-406-1	Sequence 1, Appli
	6	43.4	4.0	1083	4	US-09-252-991A-6409	Sequence 6409, Ap
	7	43.4	4.0	1104	4	US-09-252-991A-6241	Sequence 6241, Ap
c	8	41.4	3.8	7218	1	US-08-232-463-14	Sequence 14, Appl
	9	41	3.8	2437	4	US-09-833-381-1478	Sequence 1478, Ap
	10	41	3.8	2437	4	US-09-833-381-1479	Sequence 1479, Ap
	11	40	3.7	738	4	US-09-252-991A-13238	Sequence 13238, A
	12	38.2	3.5	363	4	US-09-252-991A-9714	Sequence 9714, Ap
	13	38.2	3.5	1254	4	US-09-252-991A-9674	Sequence 9674, Ap
c	14	38.2	3.5	3654	4	US-09-252-991A-9533	Sequence 9533, Ap
	15	36.4	3.4	486	4	US-09-252-991A-13452	Sequence 13452, A
	16	36.4	3.4	1209	4	US-09-252-991A-13551	Sequence 13551, A
c	17	36.4	3.4	1230	4	US-09-252-991A-13890	Sequence 13890, A
c	18	36.4	3.4	1239	4	US-09-252-991A-13787	Sequence 13787, A
	19	36	3.3	34094	4	US-09-292-034-1	Sequence 1, Appli
	20	35.6	3.3	642	4	US-09-252-991A-12925	Sequence 12925, A
	21	35.6	3.3	762	4	US-09-252-991A-13075	Sequence 13075, A
	22	35.4	3.3	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	23	35.4	3.3	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	24	35	3.2	819	4	US-09-252-991A-15630	Sequence 15630, A
	25	35	3.2	885	4	US-09-252-991A-15757	Sequence 15757, A
c	26	35	3.2	921	4	US-09-252-991A-15590	Sequence 15590, A
c	27	34.8	3.2	1935	4	US-09-252-991A-8513	Sequence 8513, Ap
	28	34.8	3.2	1944	4	US-09-252-991A-8364	Sequence 8364, Ap
	29	34.8	3.2	2073	4	US-09-252-991A-8332	Sequence 8332, Ap
	30	34.6	3.2	627	4	US-09-252-991A-11427	Sequence 11427, A
c	31	34.6	3.2	650	4	US-09-669-751-140	Sequence 140, App
	32	34.6	3.2	909	4	US-09-252-991A-11391	Sequence 11391, A
c	33	34.6	3.2	948	4	US-09-252-991A-11445	Sequence 11445, A
	34	34.4	3.2	213	4	US-09-930-181-5	Sequence 5, Appli
	35	34.4	3.2	2908	4	US-09-930-181-1	Sequence 1, Appli
	36	34.4	3.2	3364	4	US-09-930-181-3	Sequence 3, Appli
	37	34.2	3.2	690	4	US-09-252-991A-14058	Sequence 14058, A
c	38	34	3.1	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c	39	34	3.1	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	40	33.8	3.1	399	4	US-09-252-991A-5985	Sequence 5985, Ap
	41	33.8	3.1	1932	4	US-09-252-991A-5919	Sequence 5919, Ap
c	42	33.8	3.1	2214	4	US-09-252-991A-5968	Sequence 5968, Ap
	43	33.8	3.1	2370	4	US-09-489-039A-1363	Sequence 1363, Ap
	44	33.4	3.1	1756	4	US-09-166-350-30	Sequence 30, Appl
	45	33.4	3.1	1756	4	US-09-023-655-954	Sequence 954, App

ALIGNMENTS

RESULT 1
US-09-516-143A-3
; Sequence 3, Application US/09516143A
; Patent No. 6333182
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Glycosylation Enzymes
; FILE REFERENCE: PF505PCT
; CURRENT APPLICATION NUMBER: US/09/516,143A
; CURRENT FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/122,409
; PRIOR FILING DATE: 1999-03-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1077)
US-09-516-143A-3

Query Match 100.0%; Score 1080; DB 4; Length 1080;
Best Local Similarity 100.0%; Pred. No. 1.3e-302;
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
```

```
Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
```

```
RESULT 2
US-09-620-312D-1035
; Sequence 1035, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 1035
;   LENGTH: 1230
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (119)..(1198)
US-09-620-312D-1035

  Query Match             99.7%;  Score 1076.8;  DB 4;  Length 1230;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-301;
  Matches 1078;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 178

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 238

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 298

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        299 AAGTTTAATCGGAAAGCCTTGGACAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 358

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 418

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 478

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 538

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 598

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 658

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 718

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 778

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 838

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        839 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 898

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 958

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 1018

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 1078

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 GGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1138

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1198


RESULT 3
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
;  APPLICANT: Bult et al.
;  TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ
; Patent No. 6503729
;  TITLE OF INVENTION:  jannaschii
;  FILE REFERENCE: PB275
;  CURRENT APPLICATION NUMBER: US/08/916,421B
;  CURRENT FILING DATE:  1997-08-22
;  PRIOR APPLICATION NUMBER: US 60/024,428
;  PRIOR FILING DATE: 1996-08-22
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1664976
;   TYPE: DNA
;   ORGANISM: Methanococcus jannaschii
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (28222)..(28222)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (28257)..(28258)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84773)..(84773)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84808)..(84808)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84812)..(84812)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98120)..(98120)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98159)..(98159)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98239)..(98239)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98266)..(98266)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98343)..(98343)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (103998)..(103998)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (148948)..(148948)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (163385)..(163385)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (191989)..(191989)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (191995)..(191995)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (231980)..(231980)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (234187)..(234187)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (234220)..(234220)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (234814)..(234814)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (309398)..(309398)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (309418)..(309418)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (312837)..(312837)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (312993)..(312993)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (319226)..(319226)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (559167)..(559167)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (559241)..(559241)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (600992)..(600992)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (622708)..(622708)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (657081)..(657081)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (657203)..(657203)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (674435)..(674435)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (682442)..(682442)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (713652)..(713652)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (741684)..(741684)
;   OTHER INFORMATION: n equals a, t, c, or g
```

```
;   NAME/KEY: misc_feature
;   LOCATION: (779455)..(779455)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (779676)..(779676)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (855539)..(855539)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (871619)..(871619)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1084830)..(1084830)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1096846)..(1096846)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1119881)..(1119881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1130881)..(1130881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1310988)..(1310988)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1313224)..(1313224)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1349473)..(1349473)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1349491)..(1349491)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1470091)..(1470091)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1569020)..(1569020)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1602912)..(1602912)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1603734)..(1603734)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1637998)..(1637998)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1664854)..(1664855)
;   OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

  Query Match             6.8%;  Score 73.6;  DB 4;  Length 1664976;
  Best Local Similarity   46.9%;  Pred. No. 9.9e-10;
  Matches  300;  Conservative    0;  Mismatches  334;  Indels    6;  Gaps    2;

Qy        275 ATGACCAGTACAGGGAGCTGCAGAGGTACGCCGAGGAGGTTGGGATCTTCTTCACTGCCT 334
              | ||  | | |   || ||  |    || || ||  |  |||| ||  | || |   |
Db     465673 AGGAGGAATTCTATGAACTAAAAGAATATGCAGAAAAAATTGGAATTATGTTTATCTCAA 465732

Qy        335 CTGGCATGGATGAGATGGCAGTTGAATTCCTGCATGAACTGAATGTTCCATTTTTCAAAG 394
              |     | |||   |     |||||  |  |  || || ||||||| ||    || |||
Db     465733 CACCATTAGATTTAAAATATGTTGATATATTAAATAAAATGAATGTGCCTGCATTTAAAA 465792

Qy        395 TTGGATCTGGAGACACTAATAATTTTCCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCC 454
              |||  ||||| ||   |   || ||| |  | |||||   || |||| |||    |
Db     465793 TTGCCTCTGGTGATTTAACCTTTTATCCCTTATTAGAAAAAGTGGCAAAAACAGGCAAGC 465852

Qy        455 CAATGGTGATCTCCAGTGGGATGCAGTCAAT---GGACACCATGAAGCAAGTTTATCAGA 511
              |  || |  | || |  || |||      ||   ||| |    | |   |||| |
Db     465853 CGGTGATTTTATCTACAGGAATGTCTGATATTGGGGAAATTTGGGAAGCAGTTAAAGTTT 465912

Qy        512 TCGTGAAGCCCCTCAACCCCAACTTCTGCTTCTTGCAGTGTACCAGCGCATACCCGCTCC 571
              | |  ||          |    |  |    || ||||| ||||       | |||||   ||
Db     465913 TAGAAAATAATGGATGCAGGGATATTATTTTATTGCATTGTATTTCATCTTACCCAACCC 465972

Qy        572 AGCCTGAGGACGTCAACCTGCGGGTCATCTCGGAATATCAGAAGCTCTTTCCTGACATTC 631
                  ||| || |||||  |    |  ||        |   | |   | ||   |  || |
Db     465973 CTTATGAAGATGTCAATTTAAACGCTATTAAAACCTTGAAAAGTATATTCAAT---ATCC 466029

Qy        632 CCATAGGGTATTCTGGGCATGAAACAGGCATAGCGATATCTGTGGCCGCAGTGGCTCTGG 691
              |  | || |||||||  |||  |   || |||        | || |   | || ||  | |
Db     466030 CTGTGGGATATTCTGACCATACATTGGGAATACTCGCCCCAGTAGTTTCTGTTGCCTTAG 466089

Qy        692 GGGCCAAGGTGTTGGAACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACT 751
              | ||  | ||  | ||    ||| | || || || || |    | | ||   ||| ||
Db     466090 GAGCGGATGTTATTGAGAAGCACTTTACCTTAGATAAAAATATGGAAGGTCCTGATCATG 466149

Qy        752 CGGCCTCGCTGGAGCCTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGC 811
              |   ||   || || | |||| |   ||  |||      |  |  | || ||
Db     466150 CTTTGTCAGCAGACCCAGAAGAATTTAAGGAAATGGTTAATAACATAAGATTAGTTGAAA 466209

Qy        812 GTGCCCTGGGCTCCCCAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGC 871
                 || ||         |||   |  |||| | ||| |  | |   |  ||   |
Db     466210 AAATGCTTGGAAGTGGGGAAAAGATACCAATGCCTTCTGAAAGAGACGTTATTGTTGAAG 466269

Qy        872 TGGGCAAGTCTGTGGTGGCCAAAGTGAAAATTCCGGAAGG 911
                | |    | | || || |||   || ||    ||||
Db     466270 CAAGAAGAAGTATTGTAGCAAAAAGAAATATTAAAAAAGG 466309


RESULT 4
US-08-976-259-81
; Sequence 81, Application US/08976259
; Patent No. 6316609
;  GENERAL INFORMATION:
;    APPLICANT:  Dillon, Patrick J.
;    APPLICANT:  Choi, Gil H.
;    APPLICANT:  Welch, Rodney A.
;    TITLE OF INVENTION:  Nucleotide Sequence of Escherichia coli
; Patent No. 6316609
;    NUMBER OF SEQUENCES:  142
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Ave, N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/976,259
;      FILING DATE:  Herewith
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/031,626 AND US 60/061,953
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.0740002/EKS/CBM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  81:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1234 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
```

```
;      TOPOLOGY:  linear
US-08-976-259-81

  Query Match             4.9%;  Score 52.6;  DB 4;  Length 1234;
  Best Local Similarity   48.5%;  Pred. No. 2.9e-05;
  Matches  179;  Conservative    0;  Mismatches  184;  Indels    6;  Gaps    1;

Qy          62 TCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATCC 121
               | || || |||||       || |||   || ||  |||||  |   |||  |||||
Db         536 TGATCGCAGAGATTACTACTAATCACTTTGGTGATATGGACCGATTGAAGTCAATGATTG 595

Qy         122 GCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTCA 181
               |||||||  ||   |||||||| | |  |||  | ||||||  ||||  | |||
Db         596 TAGCGGCCAAGCAGGCAGGGGCTGACTATATCAAACTGCAGAAGCGTGATGTTGAAAGTT 655

Qy         182 AGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACGT 241
               | ||   || |      |||||   || |||| |||       | | |  || |  || |
Db         656 TCTATAGCAGGGAGAAGCTGGAGTCACCGTACAACTC------TCCTTTTGGCACCACCT 709

Qy         242 ACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGGT 301
               ||||  |  |    |  | || |||  | || || |       |  |    |
Db         710 TTAGGGACTATCGGCATGGCATTGAACTCAATGAAGAGCAATTTTCCTTTGTCGACTCTT 769

Qy         302 ACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAAT 361
               |    | |||  |||| |||  ||     ||| |||     | |||  |    |   ||| |
Db         770 TCTGTAAAGAGATTGGTATCGGCTGGTTTGCTTCTATTTTAGATATGCCCTCGTATGAGT 829

Qy         362 TCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTTC 421
               || | |   || |  |         |  ||||  |   |||          || |  ||
Db         830 TCATTCGGCAATTTGAACCAGATATGATCAAGCTACCATCAACTATATCTGAACATAAAG 889

Qy         422 CTTATCTGG 430
               |||| |||
Db         890 ATTATTTGG 898


RESULT 5
US-09-495-406-1
; Sequence 1, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
;  APPLICANT: Gilbert, Michel
;  APPLICANT:  Wakarchuk, Warren W.
;  APPLICANT:  National Research Council of Canada
;  TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
;  TITLE OF INVENTION:  Gangliosides and Ganglioside Mimics
;  FILE REFERENCE: 019633-000110US
;  CURRENT APPLICATION NUMBER: US/09/495,406
;  CURRENT FILING DATE:  2000-01-31
;  PRIOR APPLICATION NUMBER: US 60/118,213
;  PRIOR FILING DATE: 1999-02-01
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 11474
;   TYPE: DNA
;   ORGANISM: Campylobacter jejuni
;   FEATURE:
;   OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
;   OTHER INFORMATION: including LOS biosynthesis locus
US-09-495-406-1

  Query Match             4.1%;  Score 44.4;  DB 4;  Length 11474;
  Best Local Similarity   47.2%;  Pred. No. 0.022;
  Matches  169;  Conservative    0;  Mismatches  186;  Indels    3;  Gaps    1;

Qy         363 CCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTTCC 422
               | |  | ||  ||   |||      ||| ||  |||| || || ||   |||||||| |||
Db        7271 CTTAGAAGATATGGGAGTTAGTGCTTTTAAGATTGGTTCAGGTGAGTGTAATAATTATCC 7330

Qy         423 TTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAGTC 482
               | |  || | | |||          || ||| |  |   ||  |||||| |
Db        7331 GCTTATTAAACACATAGCAGCCTTTAAAAAGCCTATGATAGTTAGCACAGGGATGAATAG 7390

Qy         483 AATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGCTT 542
               || || |  || || | |  |     | ||| |       |     |       |||    ||
Db        7391 TATTGAAAGTATAAAACCAACTGTAAAAATCTTATTAGACAATGAAATTCCCTTTGTTTT 7450

Qy         543 CTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATCTC 602
               ||||    ||||      |||||   || || | |    || |   |   |  ||
Db        7451 AATGCACACAACCAATCTTTACCCAACCCCGCATAATCTTGTAAGATTAAACGCTATGCT 7510

Qy         603 GGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGCAT 662
               ||||   | ||    ||| ||   ||     |||| |     |  ||   |||||  |
Db        7511 TGAATTAAAAAAAGAATTTTCTTGTATGG---TAGGCTTAAGCGACCACACAACAGATAA 7567

Qy         663 AGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
               |  | |  |  || || || || || ||     ||| | ||| | ||  | |||
Db        7568 TCTTGCGTGTTTAGGTGCGGTTGCACTTGGTGCTTGTGTGCTTGAAAGACATTTTACT 7625


RESULT 6
US-09-252-991A-6409
; Sequence 6409, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 6409
;   LENGTH: 1083
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6409

  Query Match             4.0%;  Score 43.4;  DB 4;  Length 1083;
  Best Local Similarity   49.1%;  Pred. No. 0.013;
  Matches  144;  Conservative    0;  Mismatches  146;  Indels    3;  Gaps    1;

Qy          26 CCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTCATCATTGCCGAGATCGGCCAGAACC 85
               ||  |||| |  || |    |  ||   | ||     |||   |  | ||   | | |
Db         423 CCAAGCGCGGTCTGTGGCATCGCCAGGGGCGCGAGTACATCCACATCAACGCGGTGCAGC 482

Qy          86 ACCAGGGCGACCTGGACGTAGCCAAGCGCATGATCCGCATGGCCAAGGAGTGTGGGGCTG 145
               | | ||||  ||| |||  |  |  |||     |||| | | | || || ||||    |
Db         483 CCAATGGCGTGCTGTACGGCGTGACCCGC---TACCGCTTCGACGAGCAGCGTGGCCTGG 539

Qy         146 ATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTCAAGTTTAATCGGAAAGCCTTGGAGA 205
               | |  |||  |||   ||| | | ||          |    | |  || ||      |||
Db         540 AATCGGCCAGTTTCGCCAAGCGCGCGCGCTTCGAGACCGATCACTGGCAACTGGAAGAGG 599

Qy         206 GGCCATACACCTCGAAGCATTCCTGGGGGAAGACGTACGGGGAGCACAAACGACATCTGG 265
               |   ||||| |  |||| |  |||  |||   | || || |  |||  | ||    |
Db         600 TCACCACCACCTTGCTGCATCCGAGGGAAAAGCGTTCCGAGGTGGTCAAGCTACCGACCG 659

Qy         266 AGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGGTACGCCGAGGAGGTTGGG 318
               |   | |  ||| ||||  |||   ||  | |  | || | | || | | | |
Db         660 AACGCTGGGATGCCCAGCTCAGCCCGCAACTGCTGAACACTGTGGTGATGGAG 712


RESULT 7
US-09-252-991A-6241
; Sequence 6241, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
```

```
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 6241
;   LENGTH: 1104
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6241

  Query Match             4.0%;  Score 43.4;  DB 4;  Length 1104;
  Best Local Similarity   49.1%;  Pred. No. 0.013;
  Matches  144;  Conservative    0;  Mismatches  146;  Indels    3;  Gaps    1;

Qy         26 CCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTCATCATTGCCGAGATCGGCCAGAACC 85
Db        467 CCAAGCGCGGTCTGTGGCATCGCCAGGGGCGCGAGTACATCCACATCAACGCGGTGCAGC 526

Qy         86 ACCAGGGCGACCTGGACGTAGCCAAGCGCATGATCCGCATGGCCAAGGAGTGTGGGGCTG 145
Db        527 CCAATGGCGTGCTGTACGGCGTGACCCGC---TACCGCTTCGACGAGCAGCGTGGCCTGG 583

Qy        146 ATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTCAAGTTTAATCGGAAAGCCTTGGAGA 205
Db        584 AATCGGCCAGTTTCGCCAAGCGCGCGCGCTTCGAGACCGATCACTGGCAACTGGAAGAGG 643

Qy        206 GGCCATACACCTCGAAGCATTCCTGGGGGAAGACGTACGGGGAGCACAAACGACATCTGG 265
Db        644 TCACCACCACCTTGCTGCATCCGAGGGAAAAGCGTTCCGAGGTGGTCAAGCTACCGACCG 703

Qy        266 AGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGGTACGCCGAGGAGGTTGGG 318
Db        704 AACGCTGGGATGCCCAGCTCAGCCCGCAACTGCTGAACACTGTGGTGATGGAG 756


RESULT 8
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             3.8%;  Score 41.4;  DB 1;  Length 7218;
  Best Local Similarity   4.5%;  Pred. No. 0.13;
  Matches   18;  Conservative  209;  Mismatches  170;  Indels    0;  Gaps    0;

Qy         88 CAGGGCGACCTGGACGTAGCCAAGCGCATGATCCGCATGGCCAAGGAGTGTGGGGCTGAT 147
Db       1435 CRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1376

Qy        148 TGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTCAAGTTTAATCGGAAAGCCTTGGAGAGG 207
Db       1375 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1316

Qy        208 CCATACACCTCGAAGCATTCCTGGGGGAAGACGTACGGGGAGCACAAACGACATCTGGAG 267
Db       1315 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1256

Qy        268 TTCAGCCATGACCAGTACAGGGAGCTGCAGAGGTACGCCGAGGAGGTTGGGATCTTCTTC 327
Db       1255 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1196

Qy        328 ACTGCCTCTGGCATGGATGAGATGGCAGTTGAATTCCTGCATGAACTGAATGTTCCATTT 387
Db       1195 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1136

Qy        388 TTCAAAGTTGGATCTGGAGACACTAATAATTTTCCTTATCTGGAAAAGACAGCCAAAAAA 447
Db       1135 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1076

Qy        448 GGTCGCCCAATGGTGATCTCCAGTGGGATGCAGTCAA 484
Db       1075 RRRRRRRRRATCGCAAGCTCCCTCGACCTGCAGCCAA 1039


RESULT 9
US-09-833-381-1478
; Sequence 1478, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1478
;   LENGTH: 2437
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2437)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1478

  Query Match             3.8%;  Score 41;  DB 4;  Length 2437;
```

```
  Best Local Similarity   52.7%;  Pred. No. 0.095;
  Matches   89;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy        724 GACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCCGAG 783
              | ||||     |||||   |  |   |   |||     |  |||  ||| | | ||    |
Db       1803 GCCAAGCGGCAGAAGGACCGCAAGGCCATCGCCCGCTTCCAGCAGGGACAGCAGGGGCGG 1862

Qy        784 CTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTGCTG 843
              | || || |   | | | |  ||||  |  |||| | ||  ||||  |   ||||  |  |
Db       1863 CAGGAGCAGCAGGAGGGCCCAGTGGGCCCAGCCCCGAGCCGCCCAGCACTGCAGGAGAAG 1922

Qy        844 CCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCA 892
              |    ||||| ||       | ||| ||  |||  ||  ||||| || ||
Db       1923 CAGCCTGAGAAGGAGGAGGAGGAGGAGGCGGGAGAGAGTGTGGAGGACA 1971


RESULT 10
US-09-833-381-1479
; Sequence 1479, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1479
;   LENGTH: 2437
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2437)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1479

  Query Match             3.8%;  Score 41;  DB 4;  Length 2437;
  Best Local Similarity   52.7%;  Pred. No. 0.095;
  Matches   89;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy        724 GACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCCGAG 783
              | ||||     |||||   |  |   |   |||     |  |||  ||| | | ||    |
Db       1803 GCCAAGCGGCAGAAGGACCGCAAGGCCATCGCCCGCTTCCAGCAGGGACAGCAGGGGCGG 1862

Qy        784 CTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTGCTG 843
              | || || |   | | | |  ||||  |  |||| | ||  ||||  |   ||||  |  |
Db       1863 CAGGAGCAGCAGGAGGGCCCAGTGGGCCCAGCCCCGAGCCGCCCAGCACTGCAGGAGAAG 1922

Qy        844 CCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCA 892
              |    ||||| ||       | ||| ||  |||  ||  ||||| || ||
Db       1923 CAGCCTGAGAAGGAGGAGGAGGAGGAGGCGGGAGAGAGTGTGGAGGACA 1971


RESULT 11
US-09-252-991A-13238
; Sequence 13238, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 13238
;   LENGTH: 738
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13238

  Query Match             3.7%;  Score 40;  DB 4;  Length 738;
  Best Local Similarity   55.9%;  Pred. No. 0.1;
  Matches   76;  Conservative    0;  Mismatches   60;  Indels    0;  Gaps    0;

Qy        218 CGAAGCATTCCTGGGGGAAGACGTACGGGGAGCACAAACGACATCTGGAGTTCAGCCATG 277
              ||  ||  |||| |  ||| |  |||  ||| | ||| ||  | ||| |  ||      |
Db        317 CGCCGCCGTCCTTGCAGAACATCTACAAGGAACTCAACCGCGACCTGAACATCCCGATTC 376

Qy        278 ACCAGTACAGGGAGCTGCAGAGGTACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTG 337
              | |  || |  | |||||| | |  || ||| |||  | | | ||  ||| | ||||
Db        377 CCAACAACGGCTACCTGCAGCGCTGGGCGGAGCAGGGCGTGCTGTTGCTCAATACCTCCC 436

Qy        338 GCATGGATGAGATGGC 353
              |  |  |||  |||
Db        437 TGACCGTCGAGCAGGC 452


RESULT 12
US-09-252-991A-9714
; Sequence 9714, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 9714
;   LENGTH: 363
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9714

  Query Match             3.5%;  Score 38.2;  DB 4;  Length 363;
  Best Local Similarity   49.3%;  Pred. No. 0.23;
  Matches  100;  Conservative    0;  Mismatches  103;  Indels    0;  Gaps    0;

Qy        697 AAGGTGTTGGAACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCC 756
              | | ||  | | | | |    |  ||  |  || | | |    | |||  || || ||||
Db         19 ATGATGGCCGCAGGCCTCGCCGCGCTGACCCTGAGCGGTTGTGCGAGCAACGACCCGGCT 78

Qy        757 TCGCTGGAGCCTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCC 816
              |||||||||         ||| |||| | || || |      |       | || | |||||
Db         79 CCGCTGGAGCAGATGCGCCTGACCGAACAGGCGCTGGAGCAGGCCAAGGCGGTGGGTGCC 138

Qy        817 CTGGGCTCCCCAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGC 876
              |  |   |  || ||| |  ||| ||  || | |  |  |  | | || | |
Db        139 ACCGACGACGTGGCCGAGCTGAAACTGGCCCAGGACAAGTACGCCGCGGCGCAGATCGCG 198

Qy        877 AAGTCTGTGGTGGCCAAAGTGAA 899
              | | | | || | ||  |   |||
Db        199 ATGACCGCGGAGTCCTACAAGAA 221


RESULT 13
US-09-252-991A-9674
; Sequence 9674, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
```

```
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 9674
;   LENGTH: 1254
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9674

  Query Match             3.5%;  Score 38.2;  DB 4;  Length 1254;
  Best Local Similarity   49.3%;  Pred. No. 0.43;
  Matches  100;  Conservative    0;  Mismatches  103;  Indels    0;  Gaps    0;

Qy        697 AAGGTGTTGGAACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCC 756
              | | ||   | | | | |    | ||  |  || | |    | |||  || || ||||
Db        101 ATGATGGCCGCAGGCCTCGCCGCGCTGACCCTGAGCGGTTGTGCGAGCAACGACCCGGCT 160

Qy        757 TCGCTGGAGCCTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCC 816
               |||||||||        ||| |||| | || || |      |       | || | |||||
Db        161 CCGCTGGAGCAGATGCGCCTGACCGAACAGGCGCTGGAGCAGGCCAAGGCGGTGGGTGCC 220

Qy        817 CTGGGCTCCCCAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGC 876
              | |   |    || ||| |   ||| ||   || | |   |   |   | | || | |
Db        221 ACCGACGACGTGGCCGAGCTGAAACTGGCCCAGGACAAGTACGCCGCGGCGCAGATCGCG 280

Qy        877 AAGTCTGTGGTGGCCAAAGTGAA 899
              | | | | || | ||  |  |||
Db        281 ATGACCGCGGAGTCCTACAAGAA 303


RESULT 14
US-09-252-991A-9533/c
; Sequence 9533, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 9533
;   LENGTH: 3654
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9533

  Query Match             3.5%;  Score 38.2;  DB 4;  Length 3654;
  Best Local Similarity   49.3%;  Pred. No. 0.75;
  Matches  100;  Conservative    0;  Mismatches  103;  Indels    0;  Gaps    0;

Qy        697 AAGGTGTTGGAACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCC 756
              | | ||   | | | | |    |  ||  |  || | |    | |||  || || ||||
Db        550 ATGATGGCCGCAGGCCTCGCCGCGCTGACCCTGAGCGGTTGTGCGAGCAACGACCCGGCT 491

Qy        757 TCGCTGGAGCCTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCC 816
               |||||||||        ||| |||| | || || |      |       | || | |||||
Db        490 CCGCTGGAGCAGATGCGCCTGACCGAACAGGCGCTGGAGCAGGCCAAGGCGGTGGGTGCC 431

Qy        817 CTGGGCTCCCCAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGC 876
              | |   |    || ||| |   ||| ||   || | |   |   |   | | || | |
Db        430 ACCGACGACGTGGCCGAGCTGAAACTGGCCCAGGACAAGTACGCCGCGGCGCAGATCGCG 371

Qy        877 AAGTCTGTGGTGGCCAAAGTGAA 899
              | | | | || | ||  |  |||
Db        370 ATGACCGCGGAGTCCTACAAGAA 348


RESULT 15
US-09-252-991A-13452
; Sequence 13452, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 13452
;   LENGTH: 486
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13452

  Query Match             3.4%;  Score 36.4;  DB 4;  Length 486;
  Best Local Similarity   56.8%;  Pred. No. 0.89;
  Matches   67;  Conservative    0;  Mismatches   51;  Indels    0;  Gaps    0;

Qy        196 GCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACGTACGGGGAGCACAAA 255
              ||||  |   ||||  |   |||   |   | ||| |   | | | || ||||||  |
Db         74 GCCTATGTCGCGCCAGGCGAGTCGTGCCGGGCGTGGCGCCTGGCCTTCGCGGAGCAGAGC 133

Qy        256 CGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGGTACGCCGAGGAGG 313
              |  || | | ||  | || | | |||| || | || | |||||    ||||||  || |
Db        134 CATCAGCAGCAGGCCGGCGAGGGCCAGCACGGCGATCAGCAGATAGGCGCCGACGAAG 191


Search completed: September 16, 2004, 05:31:28
Job time : 692 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 16, 2004, 04:04:27 ; Search time 5258 Seconds
(without alignments)
1036.752 Million cell updates/sec

Title: US-10-759-277-3
Perfect score: 1080
Sequence: 1 atgccgctggagctggagct.....gcaaaaaaatcaagtcttaa 1080

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3327077 seqs, 2523723180 residues

Total number of hits satisfying chosen parameters: 6654154

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1080	100.0	1080	9	US-09-984-205-3	Sequence 3, Appli
	2	1080	100.0	1080	9	US-09-930-440B-5	Sequence 5, Appli
	3	1080	100.0	1080	17	US-10-759-277-3	Sequence 3, Appli
	4	1078	99.8	1268	16	US-10-264-237-1192	Sequence 1192, Ap
	5	1076.8	99.7	1230	15	US-10-037-270-1035	Sequence 1035, Ap
	6	1076.8	99.7	1230	16	US-10-117-722-1035	Sequence 1035, Ap
c	7	418	38.7	464	9	US-09-880-107-1514	Sequence 1514, Ap
	8	371.8	34.4	486	10	US-09-918-995-5220	Sequence 5220, Ap
	9	336.8	31.2	442	10	US-09-918-995-28620	Sequence 28620, A
	10	327.8	30.4	417	10	US-09-918-995-25343	Sequence 25343, A
	11	315.8	29.2	480	10	US-09-918-995-19353	Sequence 19353, A
	12	280.6	26.0	404	15	US-10-181-447A-18	Sequence 18, Appl
	13	256.2	23.7	410	15	US-10-106-698-2649	Sequence 2649, Ap
	14	180.4	16.7	203	13	US-10-085-783A-31232	Sequence 31232, A
	15	180.4	16.7	203	16	US-10-242-535A-31232	Sequence 31232, A
	16	135	12.5	933	15	US-10-156-761-3360	Sequence 3360, Ap
c	17	135	12.5	9025608	15	US-10-156-761-1	Sequence 1, Appli
	18	101.6	9.4	17276	9	US-09-870-759-83	Sequence 83, Appl
	19	101.6	9.4	17276	10	US-09-751-708A-83	Sequence 83, Appl
	20	95.2	8.8	25020	16	US-10-192-280-1	Sequence 1, Appli
c	21	85	7.9	2387	16	US-10-108-260A-1941	Sequence 1941, Ap
	22	63	5.8	1059	9	US-09-930-440B-7	Sequence 7, Appli
	23	62.8	5.8	1000	17	US-10-343-561-31	Sequence 31, Appl
	24	62	5.7	133	9	US-09-925-299-701	Sequence 701, App
	25	62	5.7	133	10	US-09-925-299-701	Sequence 701, App
	26	60	5.6	60	10	US-09-908-975-9075	Sequence 9075, Ap
	27	52.6	4.9	1234	9	US-09-956-004-81	Sequence 81, Appl
	28	52.2	4.8	65	10	US-09-908-975-25490	Sequence 25490, A
	29	49.6	4.6	408	13	US-10-321-396C-5	Sequence 5, Appli
c	30	49.6	4.6	579	13	US-10-321-396C-15	Sequence 15, Appl
	31	49.6	4.6	618	13	US-10-321-396C-3	Sequence 3, Appli
	32	49.6	4.6	650	13	US-10-321-396C-17	Sequence 17, Appl
c	33	49.6	4.6	677	13	US-10-321-396C-16	Sequence 16, Appl
	34	49.6	4.6	692	13	US-10-321-396C-18	Sequence 18, Appl
	35	49.6	4.6	828	13	US-10-321-396C-1	Sequence 1, Appli
	36	44.4	4.1	11474	9	US-09-816-028A-1	Sequence 1, Appli
	37	44.4	4.1	11474	15	US-10-303-161-1	Sequence 1, Appli
	38	44.4	4.1	11474	15	US-10-303-118-1	Sequence 1, Appli
	39	44.4	4.1	11474	15	US-10-303-128-1	Sequence 1, Appli
	40	44.4	4.1	11474	15	US-10-303-134-1	Sequence 1, Appli
	41	44.4	4.1	11474	15	US-10-303-162-1	Sequence 1, Appli
	42	41	3.8	2437	9	US-09-833-381-1478	Sequence 1478, Ap
	43	41	3.8	2437	9	US-09-833-381-1479	Sequence 1479, Ap
	44	40.2	3.7	1266	15	US-10-156-761-6436	Sequence 6436, Ap
	45	38.8	3.6	473	10	US-09-918-995-26116	Sequence 26116, A

ALIGNMENTS

RESULT 1
US-09-984-205-3
; Sequence 3, Application US/09984205
; Patent No. US20020137175A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy A. et al.
; TITLE OF INVENTION: Human Glycosylation Enzymes
; FILE REFERENCE: PF505D1
; CURRENT APPLICATION NUMBER: US/09/984,205
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: PCT/US00/05325
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/516,143
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/122,409
; PRIOR FILING DATE: 1999-03-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1077)
US-09-984-205-3

Query Match 100.0%; Score 1080; DB 9; Length 1080;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
Qy       121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
Db       121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
Qy       181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
Db       181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
Qy       241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
Db       241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
Qy       301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
Db       301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
Qy       361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
Db       361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
Qy       421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
Db       421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
Qy       481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
Db       481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
Qy       541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
Db       541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
Qy       601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
Db       601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
Qy       661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
Db       661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
Qy       721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
Db       721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
Qy       781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
Db       781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
Qy       841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
Db       841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
Qy       901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
Db       901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
Qy       961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
Db       961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
Qy      1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
Db      1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
```

RESULT 2
US-09-930-440B-5
; Sequence 5, Application US/09930440B
; Patent No. US20020142386A1
; GENERAL INFORMATION:
; APPLICANT: Betenbaugh et al.

```
;  TITLE OF INVENTION: Engineering Intracellular Sialylation Pathways
;  FILE REFERENCE: PF509P2
;  CURRENT APPLICATION NUMBER: US/09/930,440B
;  CURRENT FILING DATE:  2001-08-16
;  PRIOR APPLICATION NUMBER: 60/227,579
;  PRIOR FILING DATE: 2000-08-25
;  PRIOR APPLICATION NUMBER: 09/516,793
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/169,624
;  PRIOR FILING DATE: 1999-12-08
;  PRIOR APPLICATION NUMBER: 60/122,582
;  PRIOR FILING DATE: 1999-03-02
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 1080
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1080)
US-09-930-440B-5

  Query Match           100.0%;  Score 1080;  DB 9;  Length 1080;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1080;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
Db          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
Db         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
Db        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
Db        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
Db        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
Db        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
Db        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
Db        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
Db        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
Db        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
Db        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
```

```
Db        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
```

RESULT 3
US-10-759-277-3
; Sequence 3, Application US/10759277
; Publication No. US20040142442A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy A. et al.
; TITLE OF INVENTION: Human Glycosylation Enzymes
; FILE REFERENCE: PF505D1
; CURRENT APPLICATION NUMBER: US/10/759,277
; CURRENT FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US/09/984,205
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: PCT/US00/05325
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/516,143
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/122,409
; PRIOR FILING DATE: 1999-03-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1077)
US-10-759-277-3

```
  Query Match           100.0%;  Score 1080;  DB 17;  Length 1080;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1080;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
```

RESULT 4
US-10-264-237-1192
; Sequence 1192, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515

```
;  PRIOR FILING DATE: 2000-05-19
;  NUMBER OF SEQ ID NOS: 2876
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1192
;    LENGTH: 1268
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (5)..(5)
;    OTHER INFORMATION: n equals a,t,g, or c
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (17)..(17)
;    OTHER INFORMATION: n equals a,t,g, or c
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (23)..(23)
;    OTHER INFORMATION: n equals a,t,g, or c
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (27)..(27)
;    OTHER INFORMATION: n equals a,t,g, or c
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (31)..(31)
;    OTHER INFORMATION: n equals a,t,g, or c
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1224)..(1227)
;    OTHER INFORMATION: n equals a,t,g, or c
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1230)..(1230)
;    OTHER INFORMATION: n equals a,t,g, or c
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1257)..(1257)
;    OTHER INFORMATION: n equals a,t,g, or c
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1265)..(1265)
;    OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-1192

  Query Match             99.8%;  Score 1078;  DB 16;  Length 1268;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1078;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 180

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 240

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        241 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCYAAGTTCCAGAAGAGTGAGCTAGAATTC 300

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 360

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 420

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 480

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 540

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 600

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 660

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 720

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 780

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 840

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 900

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 960

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 1020

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 1080

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1140

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1200

RESULT 5
US-10-037-270-1035
; Sequence 1035, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Tillinghast, John
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
```

```
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/10/037,270
;  CURRENT FILING DATE:  2002-01-04
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 1035
;   LENGTH: 1230
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (119)..(1198)
US-10-037-270-1035

  Query Match            99.7%;  Score 1076.8;  DB 15;  Length 1230;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 1078;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 178

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 238

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 298

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        299 AAGTTTAATCGGAAAGCCTTGGACAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 358

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 418

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 478

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 538

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 598

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 658

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 718

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 778

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 838

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 898

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 958

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 1018

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 1078

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 GGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1138

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1198


RESULT 6
US-10-117-722-1035
; Sequence 1035, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20030219744Alel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2BCIP
;  CURRENT APPLICATION NUMBER: US/10/117,722
;  CURRENT FILING DATE:  2002-04-04
;  PRIOR APPLICATION NUMBER: 09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 1035
;   LENGTH: 1230
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (119)..(1198)
US-10-117-722-1035

  Query Match            99.7%;  Score 1076.8;  DB 16;  Length 1230;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 1078;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 178

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 238

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 298

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        299 AAGTTTAATCGGAAAGCCTTGGACAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 358
```

```
Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 418

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 478

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 538

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 598

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 658

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 718

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 778

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 838

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 898

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 958

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 1018

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 1078

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 GGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1138

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1198
```

```
RESULT 7
US-09-880-107-1514/c
; Sequence 1514, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1514
;   LENGTH: 464
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA621146
US-09-880-107-1514

  Query Match             38.7%;  Score 418;  DB 9;  Length 464;
  Best Local Similarity   98.6%;  Pred. No. 4.9e-124;
  Matches  432;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;
```

```
Qy        643 TCTGGGCATGAAACAGGCATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTG 702
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||| ||
Db        464 TCTGGGCATGAAACAGGCATAGCGATATCTGTGGCCGCAGTGGCTCTTGGGGCCAAGTTG 405

Qy        703 TTGGAACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTG 762
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        404 TTGGAACGTCACATAACTTTGGACAAGACCTGGAAGTGGAGTGACCACTCGGCCTCGCTG 345

Qy        763 GAGCCTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGC 822
              ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        344 GAGCCTGGAGAACTGGCCGAGCTGGTGCGGTCATTGCGTCTTGTGGAGCGTGCCCTGGGC 285

Qy        823 TCCCCAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 TCCCCAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCT 225

Qy        883 GTGGTGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTG 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GTGGTGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTG 165

Qy        943 AAGGTGGGTGAGCCCAAAGCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAG 1002
              |||||||||| ||||||| ||||||||||||||||||||||||||||||||||||||||
Db        164 AAGGTGGGTGA-CCCAAAGGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAG 106

Qy       1003 GTCCTGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        105 GTCCTGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGC 46

Qy       1063 AAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||
Db         45 AAAAAAATCAAGTCTTAA 28
```

```
RESULT 8
US-09-918-995-5220
; Sequence 5220, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5220
;   LENGTH: 486
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(486)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-5220
```

```
  Query Match             34.4%;  Score 371.8;  DB 10;  Length 486;
  Best Local Similarity   99.5%;  Pred. No. 4.3e-109;
  Matches  373;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 170

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        171 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGATGTAGCCAAGCGCATGATC 230

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 290

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 350

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 410

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        411 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGCTGAA 470

Qy        361 TTCCTGCATGAACTG 375
              |||||||||||||||
Db        471 TTCCTGCATGAACTG 485
```

RESULT 9
US-09-918-995-28620
; Sequence 28620, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 28620
; LENGTH: 442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(442)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-28620

```
  Query Match             31.2%;  Score 336.8;  DB 10;  Length 442;
  Best Local Similarity   98.0%;  Pred. No. 8.5e-98;
  Matches  341;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 TGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTCA 61
              ||||| |  |||||||||| ||||||||||||||||||||||||||||||||||||||| |
Db         90 TGCCGTTTTAGCTGGAGCTGAGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTGA 149

Qy         62 TCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATCC 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 TCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATCC 209

Qy        122 GCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTCA 181
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        210 GCATGGCCAAGGAGTGTGGGGCTGATTGTGCTAAGTTCCAGAAGAGTGAGCTAGAATTCA 269

Qy        182 AGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACGT 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 AGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACGT 329

Qy        242 ACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGGT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 ACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGGT 389

Qy        302 ACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGA 349
              |||||||||||| ||||||||||||||||||||||||||||||||||
Db        390 ACGCCGAGGAGGCTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGA 437
```

RESULT 10
US-09-918-995-25343
; Sequence 25343, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 25343
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(417)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-25343

```
  Query Match             30.4%;  Score 327.8;  DB 10;  Length 417;
  Best Local Similarity   99.4%;  Pred. No. 6.6e-95;
  Matches  329;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        750 CTCGGCCTCGCTGGAGCCTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGA 809
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 CTCGGCCTCGCTGGAGCCTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGA 108

Qy        810 GCGTGCCCTGGGCTCCCCAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAA 869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 GCGTGCCCTGGGCTCCCCAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAA 168

Qy        870 GCTGGGCAAGTCTGTGGTGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGA 929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 GCTGGGCAAGTCTGTGGTGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGA 228

Qy        930 CATGCTCACCGTGAAGGTGGGTGAGCCCAAAGCCTATCCTCCTGAAGACATCTTTAATCT 989
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        229 CATGCTCACCGTGAAGGTGGGTGAGCCCAAAGGCTATCCTCCTGAAGACATCTTTAATCT 288

Qy        990 AGTGGGCAAGAAGGTCCTGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAATTGGT 1049
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        289 AGTGGGCAAGAAGGTCCTGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAACTGGT 348

Qy       1050 AGATAATCATGGCAAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||
Db        349 AGATAATCATGGCAAAAAAAATCAAGTCTTAA 379
```

RESULT 11
US-09-918-995-19353
; Sequence 19353, Application US/09918995
; Publication No. US20030073623A1

```
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 19353
;   LENGTH: 480
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(480)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-19353

  Query Match             29.2%;  Score 315.8;  DB 10;  Length 480;
  Best Local Similarity   99.4%;  Pred. No. 5.4e-91;
  Matches  317;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        750 CTCGGCCTCGCTGGAGCCTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGA 809
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 CTCGGCCTCGCTGGAGCCTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGA 110

Qy        810 GCGTGCCCTGGGCTCCCCAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAA 869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 GCGTGCCCTGGGCTCCCCAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAA 170

Qy        870 GCTGGGCAAGTCTGTGGTGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGA 929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 GCTGGGCAAGTCTGTGGTGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGA 230

Qy        930 CATGCTCACCGTGAAGGTGGGTGAGCCCAAAGCCTATCCTCCTGAAGACATCTTTAATCT 989
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        231 CATGCTCACCGTGAAGGTGGGTGAGCCCAAAGGCTATCCTCCTGAAGACATCTTTAATCT 290

Qy        990 AGTGGGCAAGAAGGTCCTGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAATTGGT 1049
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        291 AGTGGGCAAGAAGGTCCTGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAACTGGT 350

Qy       1050 AGATAATCATGGCAAAAAA 1068
              ||||||||||||||||||
Db        351 AGATAATCATGGCAAAAAA 369


RESULT 12
US-10-181-447A-18
; Sequence 18, Application US/10181447A
; Publication No. US20030180738A1
; GENERAL INFORMATION:
;  APPLICANT: The No. US20030180738A1tingham Trent University
;  TITLE OF INVENTION: Cancer Associated Genes and their Products
;  FILE REFERENCE: No. US20030180738A1tingham Trent Uni
;  CURRENT APPLICATION NUMBER: US/10/181,447A
;  CURRENT FILING DATE:  2002-07-18
;  PRIOR APPLICATION NUMBER: PCT/GB/01/000188
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: GB0000993.6
;  PRIOR FILING DATE: 2000-01-18
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
;   LENGTH: 404
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(404)
;   OTHER INFORMATION: n = g, a, t, or c
US-10-181-447A-18

  Query Match             26.0%;  Score 280.6;  DB 15;  Length 404;
  Best Local Similarity   98.3%;  Pred. No. 1.2e-79;
  Matches  283;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 163

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 ATCATTGNCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 223

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 283

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        284 AAGTTTAATCGGAAAGCCTTGGACAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 343

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGG 288
              ||||||||||||||||||||||||||||||||||||||||||||||   ||
Db        344 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTCAGGG 391


RESULT 13
US-10-106-698-2649
; Sequence 2649, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
;  FILE REFERENCE: PA005P1
;  CURRENT APPLICATION NUMBER: US/10/106,698
;  CURRENT FILING DATE:  2002-03-27
;  PRIOR APPLICATION NUMBER: PCT/US00/26524
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US 60/157,137
;  PRIOR FILING DATE: 1999-09-29
;  PRIOR APPLICATION NUMBER: US 60/163,280
;  PRIOR FILING DATE: 1999-11-03
;  NUMBER OF SEQ ID NOS: 8564
;  SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 2649
;   LENGTH: 410
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (287)..(287)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (300)..(300)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (370)..(370)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (374)..(374)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (402)..(402)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (405)..(405)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (408)..(408)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-2649
```

```
  Query Match             23.7%;  Score 256.2;  DB 15;  Length 410;
  Best Local Similarity   97.2%;  Pred. No. 9.2e-72;
  Matches  280;  Conservative    0;  Mismatches    6;  Indels    2;  Gaps    2;

Qy        548 AGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATCTCGGAAT 607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 AGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATCTCGGAAT 169

Qy        608 ATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGCATAGCGA 667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 ATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGCATAGCGA 229

Qy        668 TATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACTTTGGACA 727
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        230 TATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACTTTGGNCA 289

Qy        728 AGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCCG-AGCTG 786
              |||||||||| |||||||||||||||| |||||||||||||||||||||| || |||||
Db        290 AGACCTGGAANGGGAGTGACCACTCGGACTCGCTGGAGCCTGGAGAACTGGGCGAAGCTG 349

Qy        787 GTGCGGTCAGTGCGTCTTGT-GGAGCGTGCCCTGGGCTCCCCAACCAA 833
              |||||||||||||||||| ||| ||||||||||||||||||| |||
Db        350 GTGCGGTCAGTGCGTCTTGTNGGANCGTGCCCTGGGCTCCCCAAGCAA 397


RESULT 14
US-10-085-783A-31232
; Sequence 31232, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31232
;   LENGTH: 203
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (6)..(6)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (31)..(31)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-31232

  Query Match             16.7%;  Score 180.4;  DB 13;  Length 203;
  Best Local Similarity   98.4%;  Pred. No. 1.9e-47;
  Matches  181;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        897 GAAAATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCC 956
              |||||.|||||||||||||||||||||| ||||||||||||||||||||||||||||
Db          1 GAAAANTCCGGAAGGCACCATTCTAACAATNGACATGCTCACCGTGAAGGTGGGTGAGCC 60

Qy        957 CAAAGCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGT 1016
              |||||.|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAAGGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGT 120

Qy       1017 TGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTC 1076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTC 180

Qy       1077 TTAA 1080
              ||||
Db        181 TTAA 184


RESULT 15
US-10-242-535A-31232
; Sequence 31232, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31232
;   LENGTH: 203
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (6)..(6)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (31)..(31)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-31232

  Query Match             16.7%;  Score 180.4;  DB 16;  Length 203;
  Best Local Similarity   98.4%;  Pred. No. 1.9e-47;
  Matches  181;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        897 GAAAATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCC 956
              |||||.|||||||||||||||||||||| ||||||||||||||||||||||||||||
Db          1 GAAAANTCCGGAAGGCACCATTCTAACAATNGACATGCTCACCGTGAAGGTGGGTGAGCC 60

Qy        957 CAAAGCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGT 1016
              |||||.|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAAGGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGT 120

Qy       1017 TGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTC 1076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTC 180

Qy       1077 TTAA 1080
              ||||
Db        181 TTAA 184


Search completed: September 16, 2004, 09:36:47
Job time : 5264 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 16, 2004, 01:01:16 ; Search time 4417 Seconds
(without alignments)
7301.597 Million cell updates/sec

Title: US-10-759-277-3
Perfect score: 1080
Sequence: 1 atgccgctggagctggagct.....gcaaaaaaatcaagtcttaa 1080

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1076.8	99.7	1080	29	AY414837	AY414837 Homo sapi
c	2	952.6	88.2	984	9	AL578658	AL578658 AL578658
c	3	949.6	87.9	1201	13	BX384733	BX384733 BX384733
c	4	917.8	85.0	990	9	AL578579	AL578579 AL578579
	5	883.2	81.8	1080	29	AY414839	AY414839 Mus muscu
	6	883.2	81.8	1906	11	AK076290	AK076290 Mus muscu
	7	870.6	80.6	2109	11	AK088912	AK088912 Mus muscu
	8	856.6	79.3	925	12	BI517350	BI517350 603041586
	9	856.4	79.3	987	9	AL556387	AL556387 AL556387
	10	843	78.1	1005	9	AL556278	AL556278 AL556278
	11	840	77.8	1201	9	AL515470	AL515470 AL515470
c	12	830	76.9	989	9	AL526241	AL526241 AL526241
	13	823.4	76.2	889	13	BU187658	BU187658 AGENCOURT
	14	803.2	74.4	872	13	BQ432284	BQ432284 AGENCOURT
	15	796.4	73.7	1031	13	BQ064521	BQ064521 AGENCOURT
c	16	795.8	73.7	842	9	AL522851	AL522851 AL522851
	17	792.2	73.4	1049	12	BM541313	BM541313 AGENCOURT
	18	789.4	73.1	948	29	AY414838	AY414838 Pan trogl
	19	786	72.8	1046	9	AL526281	AL526281 AL526281
	20	784.4	72.6	890	13	BU181010	BU181010 AGENCOURT
	21	775	71.8	979	12	BG681730	BG681730 602627851
	22	772	71.5	905	13	BQ893527	BQ893527 AGENCOURT
	23	764.4	70.8	1097	12	BM806099	BM806099 AGENCOURT
	24	757.4	70.1	878	9	AL522852	AL522852 AL522852
	25	754.4	69.9	878	12	BI824205	BI824205 603040573
	26	749.2	69.4	1070	9	AL544111	AL544111 AL544111
	27	737.6	68.3	1050	12	BM563430	BM563430 AGENCOURT
c	28	736.2	68.2	773	12	BI517963	BI517963 603041586
	29	735.4	68.1	845	13	BU528624	BU528624 AGENCOURT
	30	726.2	67.2	905	13	BQ649507	BQ649507 AGENCOURT
	31	722.2	66.9	883	12	BI258475	BI258475 602972467
	32	721.6	66.8	777	12	BG568099	BG568099 602586980
	33	717.4	66.4	978	12	BG114903	BG114903 602315105
	34	710.6	65.8	919	13	BU195376	BU195376 AGENCOURT
	35	710.2	65.8	813	9	AU132828	AU132828 AU132828
c	36	703.2	65.1	1031	9	AL544197	AL544197 AL544197
	37	695	64.4	706	12	BG746292	BG746292 602703522
	38	690	63.9	690	12	BG746329	BG746329 602703579
c	39	689.4	63.8	835	10	BE614630	BE614630 601504568
	40	687.8	63.7	948	10	BE793764	BE793764 601589915
c	41	686.4	63.6	769	13	BU618375	BU618375 UI-H-FH1-
c	42	682.2	63.2	779	14	CA417480	CA417480 UI-H-FE0-
	43	680.4	63.0	931	12	BI525002	BI525002 602924077
	44	678.4	62.8	844	13	BU167760	BU167760 AGENCOURT
	45	675.2	62.5	877	12	BI905036	BI905036 603169320

ALIGNMENTS

RESULT 1
AY414837
LOCUS AY414837 1080 bp DNA linear GSS 17-DEC-2003
DEFINITION Homo sapiens NANS gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY414837
VERSION AY414837.1 GI:39770796
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1080)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1080)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission

```
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .1080
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1. .>1080
                     /gene="NANS"
                     /locus_tag="HCM5357"
ORIGIN

  Query Match             99.7%;  Score 1076.8;  DB 29;  Length 1080;
  Best Local Similarity   99.8%;  Pred. No. 3.4e-268;
  Matches 1078;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCTAAGTTCCAGAAGAGTGAGCTAGAATTC 180

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080


RESULT 2
AL578658/c
LOCUS       AL578658                984 bp    mRNA    linear   EST 01-JUN-2003
DEFINITION  AL578658 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DK004YI14 3-PRIME, mRNA sequence.
ACCESSION   AL578658
VERSION     AL578658.2  GI:31316853
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 984)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 16, 2001 this sequence version replaced gi:12942945.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 5047.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DK004BE07NP1&cluster=5047.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DK004BE07NP1.
FEATURES             Location/Qualifiers
     source          1. .984
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DK004YI14"
                     /cell_type="HELA CELLS COT 25-NORMALIZED"
                     /cell_line="HELA"
                     /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match             88.2%;  Score 952.6;  DB 9;  Length 984;
  Best Local Similarity   98.4%;  Pred. No. 5.7e-236;
  Matches  966;  Conservative    3;  Mismatches   12;  Indels    1;  Gaps    1;

Qy         99 GGACGTAGCCAAGCGCATGATCCGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTT 158
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        984 GGATGTAGCCAAGCGCATGATCCGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTT 925

Qy        159 CCAGAAGAGTGAGCTAGAATTCAAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTC 218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CCAGAAGAGTGAGCTAGAATTCAAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTC 865
```

```
Qy        219 GAAGCATTCCTGGGGGAAGACGTACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGA 278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 GAAGCATTCCTGGGGGAAGACGTACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGA 805

Qy        279 CCAGTACAGGGAGCTGCAGAGGTACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGG 338
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 CCAGTACAGGGAGCTGCAGAGGTACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGG 745

Qy        339 CATGGATGAGATGGCAGTTGAATTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGG 398
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 CATGGATGAGATGGCAGTTGAATTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGG 685

Qy        399 ATCTGGAGACACTAATAATTTTCCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAAT 458
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 ATCTGGAGACACTAATAATTTTCCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAAT 625

Qy        459 GGTGATCTCCAGTGGGATGCAGTCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAA 518
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 GGTGATCTCCAGTGGGATGCAGTCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAA 565

Qy        519 GCCCCTCAACCCCAACTTCTGCTTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGA 578
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 GCCCCTCAACCCCAACTTCTGCTTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGA 505

Qy        579 GGACGTCAACCTGCGGGTCATCTCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGG 638
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGACGTCAA-CTGCGGGTCATCTCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGG 446

Qy        639 GTATTCTGGGCATGAAACAGGCATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAA 698
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 GTATTCTGGGCATGAAACAGGCATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAA 386

Qy        699 GGTGTTGGAACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTC 758
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 GGTGTTGGAACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTC 326

Qy        759 GCTGGAGCCTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCT 818
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GCTGGAGCCTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCT 266

Qy        819 GGGCTCCCCAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAA 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 GGGCTCCCCAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAA 206

Qy        879 GTCTGTGGTGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAACAATGGACATGCTCAC 938
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        205 GTCTGTGGTGGCCAAAGTGAAAATTCCGGAAGGCRCCATTCTAACAATGGACATGCTCAC 146

Qy        939 CGTGAAGGTGGGTGAGCCCAAAGCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAA 998
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        145 CGTGAAGGTGGGTGAGCCCAAAGGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAA 86

Qy        999 GAAGGTCCTGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCA 1058
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GAAGGTCCTGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAATTGGTAGATAATCA 26

Qy       1059 TGGCAAAAAAAATCAAGTCTTAA 1080
              ||  ||||||     | : |:
Db         25 TGCAAAAAAATCNNNNTNYMAM 4
```

```
RESULT 3
BX384733/c
LOCUS       BX384733               1201 bp    mRNA    linear   EST 08-MAY-2003
DEFINITION  BX384733 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
            Homo sapiens cDNA clone CS0DL001YG22 3-PRIME, mRNA sequence.
ACCESSION   BX384733
VERSION     BX384733.1  GI:30449274
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 5047.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DL001BD11NP1&cluster=5047.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DL001BD11NP1.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DL001YG22"
                     /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
                     /cell_line="RAMOS CELL LINE"
                     /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
                     25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
```

```
  Query Match            87.9%;  Score 949.6;  DB 13;  Length 1201;
  Best Local Similarity  97.1%;  Pred. No. 3.8e-235;
  Matches 1039;  Conservative   11;  Mismatches   10;  Indels   10;  Gaps    8;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||::|    |:||||||:||||||||   ||  ||||||||||||||||||
Db       1060 ATGCCGCTGGAGYKG--AGTKTGTCCCSGGCGCTGG--TGGGCGGCAACACCCGTGCTTC 1005

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              |||| ||:|||||||| |||||||||||||| ||||||||||||||:|||||||||||||||
Db       1004 ATCA-TGSCGAGATC-GCCAGAACCACCA-GGCGACCTGGACGTARCCAAGCGCATGATC 948

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              |||||||||||||||||| |||||||||| |||||||||||||||||||||||||||||||
Db        947 CGCATGGCCAAGGAGTGTG-GGCTGATTGTGCTAAGTTCCAGAAGAGTGAGCTAGAATTC 889

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        888 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCT-GGGGAAGACG 830

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 770

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              |||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        769 TACGCCGAGGAGGTTGGGATCTTCTTCAACTCCTCTGGCATGGATGAGATGGCAGTTGAA 710

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 650

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 590

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 530
```

```
Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        529 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAA-CTGCGGGTCATC 471

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||:||||||||||||||||||||||||||||:||||||
Db        470 TCGGAATATCAGAAGCTCTTYCCTGACATTCCCATAGGGTATTCTGGGCATGAMACAGGC 411

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        410 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 351

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 291

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||:||||||||||||||||||:|||:||||||||||||
Db        290 GAGCTGGTGCGGTCAGTGCKTCTTGTGGAGCGTGCCCTGGGCTCYCCAWCCAAGCAGCTG 231

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 171

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 111

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 GGCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 51

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAAT 1070
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         50 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAAT 1
```

```
RESULT 4
AL578579/c
LOCUS       AL578579                990 bp    mRNA    linear   EST 01-JUN-2003
DEFINITION  AL578579 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DK001YM22 3-PRIME, mRNA sequence.
ACCESSION   AL578579
VERSION     AL578579.1  GI:12942789
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 990)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 5047.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DK001BG11NP1&cluster=5047.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DK001BG11NP1.
FEATURES             Location/Qualifiers
     source          1. .990
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DK001YM22"
                     /cell_type="HELA CELLS COT 25-NORMALIZED"
                     /cell_line="HELA"
                     /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
```

```
  Query Match             85.0%;  Score 917.8;  DB 9;  Length 990;
  Best Local Similarity   96.6%;  Pred. No. 6.2e-227;
  Matches  958;  Conservative    8;  Mismatches   23;  Indels    3;  Gaps    3;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              |||||||||| ||||||||| ||||||||||||||||||||||||||||||||||||||
Db        990 ATGCCGCTGGGGCTGGAGCTTTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 931

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              || |||||||||||||||||||||||||||||||||||||| ||||||||||||||| |||
Db        930 ATAATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGATGTAGCCAAGCGCATAATC 871

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        870 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 811

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        810 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 751

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              |||||||||||||||||| |||||||||||||||||||||:||||||:||||||||||
Db        750 TACGGGGAGCACAAACGACTTCTGGAGTTCAGCCATGACCAGWACAGGGRGCTGCAGAGG 691

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              |||||||||||||||||||||||| ||| ||||||||||||||||||||||||||| |
Db        690 TACGCCGAGGAGGTTGGGATCTTCTTAACT-CCTCTGGCATGGATGAGATGGCAGTTGGA 632

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 572

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        571 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGGTCTCCAGTGGGATGCAG 512

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db        511 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCGACTTCTGC 452

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              |||||||||||||||||||| |||:|||||||||||||||||||| ||||||||||||
Db        451 TTCTTGCAGTGTACCAGCGCANACCSGCTCCAGCCTGAGGACGTCAA-CTGCGGGTCATC 393

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              |||| || |||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        392 TCGGNATNTCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAGACAGGC 333

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              |||||||| |||||||||| ||||||||||||||||||||||||||||| |||| |||
Db        332 ATAGCGATGTCTGTGGCCGGAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTGACATGACT 273

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              |||||||||||||||||||||||||||||||||||:|||||||||||||| |||||||
Db        272 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTSGCTGGAGCCTGGAGGACTGGCC 213

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              | |||||||||||||||||||||||||:|||||||||||||||||:|||:|| ||||
Db        212 GGGCTGGTGCGGTCAGTGCGTCTTGTGGRGCGTGCCCTGGGCTCCCCARCCARGCGGCTG 153

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 93

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGC-CCAA 959
              |||||||||||||||||||:|||||| |||||||||||||||||||||||||| ||||
```

```
Db         92 ATTCCGGAAGGCACCATTCTMACAATGAACATGCTCACCGTGAAGGTGGGTGAGCGCCAA 33

Qy        960 AGCCTATCCTCCTGAAGACATCTTTAATCTAG 991
              || ||||||||||||||||||||||||||||||
Db         32 AGGCTATCCTCCTGAAGACATCTTTAATCTAG 1


RESULT 5
AY414839
LOCUS       AY414839               1080 bp    DNA     linear   GSS 17-DEC-2003
DEFINITION  Mus musculus NANS gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY414839
VERSION     AY414839.1  GI:39770798
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1080)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 1080)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .1080
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1. .>1080
                     /gene="NANS"
                     /locus_tag="HCM5357"
ORIGIN

  Query Match           81.8%;  Score 883.2;  DB 29;  Length 1080;
  Best Local Similarity  88.6%;  Pred. No. 6.3e-218;
  Matches  957;  Conservative    0;  Mismatches  123;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              |||||||||| |||||||||||||||||||||||||||| ||  | ||||||||||||
Db          1 ATGCCGCTGGAACTGGAGCTGTGTCCCGGGCGCTGGGTGGGTGGAAAGCACCCGTGCTTC 60

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||| || |||||||||||||||||||| || ||| | || || ||||||||||||||||
Db         61 ATCATCGCGGAGATCGGCCAGAACCACCAAGGAGACATAGATGTGGCCAAGCGCATGATC 120

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||  |||||||||||||||||| || || || ||||| |||||||| ||| | || |||
Db        121 CGCACTGCCAAGGAGTGTGGGGCCGACTGCGCTAAGTTTCAGAAGAGCGAGTTGGAGTTC 180

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||| ||||| ||| ||||||| ||||| || |||||||||| ||||||||||||
Db        181 AAGTTTAACCGGAAGGCCCTGGAGAGACCATATACTTCGAAGCATTCATGGGGGAAGACG 240

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              || |||||||||| || |||||||| |||||||| |||||||||| ||||||||||||
Db        241 TATGGGGAGCACAAGCGGCATCTGGAATTCAGCCACGACCAGTACAAGGAGCTGCAGAGC 300

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
```

```
              || ||  ||||| | || ||||||||||||||||||||||||| ||||||||||||||
Db        301 TATGCGCAGGAGATCGGCATCTTCTTCACTGCCTCTGGCATGGACGAGATGGCAGTTGAG 360

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              || ||||| |||||||||||||| |||||||||||||||||||| |||||||| ||||||
Db        361 TTTCTGCACGAACTGAATGTTCCCTTTTTCAAAGTTGGATCTGGGGACACTAACAATTTT 420

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              || || ||||||||||||||||||| |||||||| || |||||||||||||| |||||||||
Db        421 CCCTACCTGGAAAAGACAGCCAAGAAAGGTCGTCCTATGGTGATCTCCAGCGGGATGCAG 480

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              |||||||||||||||||||||| |||||||||||||||||| || || ||||||||||||
Db        481 TCAATGGACACCATGAAGCAAGTCTATCAGATCGTGAAGCCGCTGAATCCCAACTTCTGC 540

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||| | || || |||||||| ||||| || ||||| ||||| | |||||||||| ||||||
Db        541 TTCCTCCAATGCACCAGCGCGTACCCACTACAGCCCGAGGATGCCAACCTGCGCGTCATC 600

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              |||||||| |||||||||||||| |||||||||| ||||||| ||||| || || |||
Db        601 TCGGAATACCAGAAGCTCTTTCCCGACATTCCCATCGGGTATTCCGGGCACGAGACGGGC 660

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              || || |||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        661 ATCGCCATATCTGTGGCCGCCGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACG 720

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              |||||||||||||||||||||||||||||| |||||||||||||||| || |||||
Db        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCAGCCTCGCTGGAGCCTGGGGAGCTGGCA 780

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              |||||||||||||| ||||| || |||||||| ||||||||||||||||||||||||||||
Db        781 GAGCTGGTGCGGTCTGTGCGCCTGGTGGAGCGGGCCCTGGGCTCCCCAACCAAGCAGCTG 840

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||
Db        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTCGGCAAGTCTGTGGTAGCCAAAGTGAAA 900

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              || || | |||||||| || ||  |||||||||||| |||||||||| |||||||||
Db        901 ATCCCAGCAGGCACCACCCTGACCCTGGACATGCTCACTGTGAAGGTGGGGGAGCCCAAA 960

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | |||||||||||||||||||| || |||| |||||| ||||| |||||||||| | |||
Db        961 GGCTATCCTCCTGAAGACATCTTCAACCTAGCGGGCAAAAAGGTGCTGGTCACTATCGAA 1020

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              || ||||||||  |||||||||||||  || || | |||  |||| ||||||||| |||||
Db       1021 GAAGATGACACGGTCATGGAAGAATCCGTGGAAAGTCACAGCAAGAAAATCAAGGCTTAA 1080


RESULT 6
AK076290
LOCUS       AK076290               1906 bp    mRNA    linear   HTC 20-SEP-2003
DEFINITION  Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched
            library, clone:4632418E04 product:sialic acid synthase, full insert
            sequence.
ACCESSION   AK076290
VERSION     AK076290.1  GI:26345279
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
```

```
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 1906)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
            Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
            Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
            Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
            Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
            Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A.,
            Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1. .1906
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:4632418E04"
                     /db_xref="MGI:1903268"
                     /db_xref="taxon:10090"
                     /clone="4632418E04"
                     /tissue_type="skin"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="0 day neonate"
     CDS             32. .1111
                     /note="unnamed protein product; putative
                     sialic acid synthase (MGD|MGI:2149820, GB|NM_053179,
                     evidence: BLASTN, 100%, match=1140)"
                     /codon_start=1
                     /protein_id="BAC36290.1"
                     /db_xref="GI:26345280"
                     /translation="MPLELELCPGRWVGGKHPCFIIAEIGQNHQGDIDVAKRMIRTAK
                     ECGADCAKFQKSELEFKFNRKALERPYTSKHSWGKTYGEHKRHLEFSHDQYKELQSYA
                     QEIGIFFTASGMDEMAVEFLHELNVPFFKVGSGDTNNFPYLEKTAKKGRPMVISSGMQ
                     SMDTMKQVYQIVKPLNPNFCFLQCTSAYPLQPEDANLRVISEYQKLFPDIPIGYSGHE
                     TGIAISVAAVALGAKVLERHITLDKTWKGSDHSASLEPGELAELVRSVRLVERALGSP
                     TKQLLPCEMACNEKLGKSVVAKVKIPAGTTLTLDMLTVKVGEPKGYPPEDIFNLAGKK
                     VLVTIEEDDTVMEESVESHSKKIKA"
ORIGIN

  Query Match             81.8%;  Score 883.2;  DB 11;  Length 1906;
  Best Local Similarity   88.6%;  Pred. No. 8.5e-218;
  Matches  957;  Conservative    0;  Mismatches  123;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              |||||||||| |||||||||||||||||||||||||||||| ||  | ||||||||||||
Db         32 ATGCCGCTGGAACTGGAGCTGTGTCCCGGGCGCTGGGTGGGTGGAAAGCACCCGTGCTTC 91

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||| || |||||||||||||||||||| || ||| | || || ||||||||||||||||
Db         92 ATCATCGCGGAGATCGGCCAGAACCACCAAGGAGACATAGATGTGGCCAAGCGCATGATC 151

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||  |||||||||||||||||| || || || ||||| |||||||| ||| | || |||
Db        152 CGCACTGCCAAGGAGTGTGGGGCCGACTGCGCTAAGTTTCAGAAGAGCGAGTTGGAGTTC 211

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||| ||||| ||| ||||||| ||||| || |||||||||| ||||||||||||
Db        212 AAGTTTAACCGGAAGGCCCTGGAGAGACCATATACTTCGAAGCATTCATGGGGGAAGACG 271

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              || |||||||||| || |||||||| |||||||| |||||||||| ||||||||||||
Db        272 TATGGGGAGCACAAGCGGCATCTGGAATTCAGCCACGACCAGTACAAGGAGCTGCAGAGC 331

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              || ||  ||||| | || ||||||||||||||||||||||||| ||||||||||||||||
Db        332 TATGCGCAGGAGATCGGCATCTTCTTCACTGCCTCTGGCATGGACGAGATGGCAGTTGAG 391

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              || ||||| |||||||||||||| ||||||||||||||||| ||||||| ||||||||||
Db        392 TTTCTGCACGAACTGAATGTTCCCTTTTTCAAAGTTGGATCTGGGGACACTAACAATTTT 451

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              || || |||||||||||||||| |||||||| || |||||||||||||| |||||||||
Db        452 CCCTACCTGGAAAAGACAGCCAAGAAAGGTCGTCCTATGGTGATCTCCAGCGGGATGCAG 511

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              |||||||||||||||||||||||| |||||||||||||||| || || ||||||||||||
Db        512 TCAATGGACACCATGAAGCAAGTCTATCAGATCGTGAAGCCGCTGAATCCCAACTTCTGC 571

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||| | || || |||||||| ||||| || ||||| ||||| | |||||||| ||||||
Db        572 TTCCTCCAATGCACCAGCGCGTACCCACTACAGCCCGAGGATGCCAACCTGCGCGTCATC 631

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              |||||||| |||||||||||||| |||||||||| |||||||| ||||| || || |||
Db        632 TCGGAATACCAGAAGCTCTTTCCCGACATTCCCATCGGGTATTCCGGGCACGAGACGGGC 691

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              || || |||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        692 ATCGCCATATCTGTGGCCGCCGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACG 751

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              |||||||||||||||||||||||||||||| |||||||||||||||||||| || |||||
Db        752 TTGGACAAGACCTGGAAGGGGAGTGACCACTCAGCCTCGCTGGAGCCTGGGGAGCTGGCA 811

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              |||||||||||||| ||||| || |||||||| ||||||||||||||||||||||||||||
Db        812 GAGCTGGTGCGGTCTGTGCGCCTGGTGGAGCGGGCCCTGGGCTCCCCAACCAAGCAGCTG 871
```

```
Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
              ||||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||
Db        872 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTCGGCAAGTCTGTGGTAGCCAAAGTGAAA 931

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960
              || || | ||||||||  || ||  |||||||||||||| |||||||||| |||||||||
Db        932 ATCCCAGCAGGCACCACCCTGACCCTGGACATGCTCACTGTGAAGGTGGGGGAGCCCAAA 991

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAA 1020
              | |||||||||||||||||||||| || |||| |||||| |||| |||||||||| | |||
Db        992 GGCTATCCTCCTGAAGACATCTTCAACCTAGCGGGCAAAAAGGTGCTGGTCACTATCGAA 1051

Qy       1021 GAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              || ||||||||  |||||||||||  || || | |||  |||| |||||||| |||||
Db       1052 GAAGATGACACGGTCATGGAAGAATCCGTGGAAAGTCACAGCAAGAAAATCAAGGCTTAA 1111
```

RESULT 7
AK088912

```
LOCUS       AK088912               2109 bp    mRNA    linear   HTC 20-SEP-2003
DEFINITION  Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN
            full-length enriched library, clone:E430030P04 product:sialic acid
            synthase, full insert sequence.
ACCESSION   AK088912
VERSION     AK088912.1  GI:26104987
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 2109)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
            Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
            Trust/MRC building Addenbrookes Hospital Cambridge) whose
            assistance we gratefully acknowledge.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1. .2109
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="NOD"
                     /db_xref="FANTOM_DB:E430030P04"
                     /db_xref="MGI:2427809"
                     /db_xref="taxon:10090"
                     /clone="E430030P04"
                     /cell_type="thymic cells"
                     /tissue_type="thymus"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="2 days neonate"
     misc_feature    41. .1121
                     /note="putative
                     sialic acid synthase (MGD|MGI:2149820, GB|NM_053179,
                     evidence: BLASTN, 100%, match=1140)"
ORIGIN
```

Query Match 80.6%; Score 870.6; DB 11; Length 2109;
Best Local Similarity 88.4%; Pred. No. 1.7e-214;
Matches 956; Conservative 0; Mismatches 124; Indels 1; Gaps 1;

```
Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              |||||||||| |||||||||||||||||||||||||||||| ||  | ||||||||||||
Db         41 ATGCCGCTGGAACTGGAGCTGTGTCCCGGGCGCTGGGTGGGTGGAAAGCACCCGTGCTTC 100

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||| || |||||||||||||||||||||| || ||| | || || ||||||||||||||||
Db        101 ATCATCGCGGAGATCGGCCAGAACCACCAAGGAGACATAGATGTGGCCAAGCGCATGATC 160

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||  |||||||||||||||||| || || || ||||| |||||||| ||| | || |||
Db        161 CGCACTGCCAAGGAGTGTGGGGCCGACTGCGCTAAGTTTCAGAAGAGCGAGTTGGAGTTC 220

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              |||||||| ||||| ||| ||||||| ||||| || |||||||||| ||||||||||||
Db        221 AAGTTTAACCGGAAGGCCCTGGAGAGACCATATACTTCGAAGCATTCATGGGGGAAGACG 280

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              || |||||||||| || |||||||| |||||||| |||||||||| ||||||||||||
Db        281 TATGGGGAGCACAAGCGGCATCTGGAATTCAGCCACGACCAGTACAAGGAGCTGCAGAGC 340

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
```

```
Db        341 TATGCGCAGGAGATCGGCATCTTCTTCACTGCCTCTGGCATGGACGAGATGGCAGTTGAG 400

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420

Db        401 TTTCTGCACGAACTGAATGTTCCCTTTTTCAAAGTTGGATCTGGGGACACTAACAATTTT 460

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480

Db        461 CCCTACCTGGAAAAGACAGCCAAGAAAGGTCGTCCTATGGTGATCTCCAGCGGGATGCAG 520

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540

Db        521 TCAATGGACACCATGAAGCAAGTCTATCAGATCGTGAAGCCGCTGAATCCCAACTTCTGC 580

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600

Db        581 TTCCTCCAATGCACCAGCGCGTACCCACTACAGCCCGAGGATGCCAACCTGCGCGTCATC 640

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660

Db        641 TCGGAATACCAGAAGCTCTTTCCCGACATTCCCATCGGGTATTCCGGGCACGAGACGGGC 700

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720

Db        701 ATCGCCATATCTGTGGCCGCCGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACG 760

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780

Db        761 TTGGACAAGACCTGGAAGGGGAGTGACCACTCAGCCTCGCTGGAGCCTGGGGAGCTGGCA 820

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840

Db        821 GAGCTGGTGCGGTCTGTGCGCCTGGTGGAGCGGGCCCTGGGCTCCCCAACCAAGCAGCTG 880

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900

Db        881 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTCGGCAAGTCTGTGGTAGCCAAAGTGAAA 940

Qy        901 ATTCCGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA 960

Db        941 ATCCCAGCAGGCACCACCCTGACCCTGGACATGCTCACTGTGAAGGTGGGGGAGCCCAAA 1000

Qy        961 GCCTATCCTCCTGAAGACATCTTTAATCTAGTGGGC-AAGAAGGTCCTGGTCACTGTTGA 1019

Db       1001 GGCTATCCTCCTGAAGACATCTTCAACCTAGCGGACAAAAAAGGTGCTGGTCACTATCGA 1060

Qy       1020 AGAGGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTA 1079

Db       1061 AGAAGATGACACGGTCATGGAAGAATCCGTGGAAAGTCACAGCAAGAAAATCAAGGCTTA 1120

Qy       1080 A 1080

Db       1121 A 1121


RESULT 8
BI517350
LOCUS       BI517350                925 bp    mRNA    linear   EST 29-AUG-2001
DEFINITION  603041586F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5163020 5',
            mRNA sequence.
ACCESSION   BI517350
VERSION     BI517350.1  GI:15342142
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 925)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11404  row: k  column: 21
            High quality sequence stop: 856.
FEATURES             Location/Qualifiers
     source          1. .925
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5163020"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_116"
                     /note="Organ: pooled colon, kidney, stomach; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 3 colons, age 26 yo male, 49 yo
                     female, 71 yo male colon; 46 yo male kidney, and pool of 2
                     stomachs, 62 yo male and 70 yo female.  Library is
                     oligo-dT primed and  directionally cloned (EcoRV site is
                     destroyed upon cloning).  Average insert size 1.4 kb,
                     insert size range 1-3 kb.  Library is normalized and
                     enriched for full-length clones and was constructed by C.
                     Gruber  (Invitrogen).  Research Genetics tracking code
                     023. Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             79.3%;  Score 856.6;  DB 12;  Length 925;
  Best Local Similarity   98.8%;  Pred. No. 4.7e-211;
  Matches  884;  Conservative    0;  Mismatches    9;  Indels    2;  Gaps    2;

Qy         188 ATCGGAAAGCCTTGGAG-AGGCCATACACCTCGAAGCATTCCTGGGGGAAGACGTACGGG 246

Db           1 ATCGGAAAGCCTTGGAGAAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACGTACGGG 60

Qy         247 GAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGGTACGCC 306

Db          61 GAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGGTACGCC 120

Qy         307 GAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAATTCCTG 366

Db         121 GAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAATTCCTG 180

Qy         367 CATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTTCCTTAT 426

Db         181 CATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTTCCTTAT 240

Qy         427 CTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAGTCAATG 486

Db         241 CTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAGTCAATG 300

Qy         487 GACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGCTTCTTG 546

Db         301 GACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGCTTCTTG 360

Qy         547 CAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATCTCGGAA 606

Db         361 CAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATCTCGGAA 420

Qy         607 TATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGCATAGCG 666

Db         421 TATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGCATAGCG 480

Qy         667 ATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACTTTGGAC 726

Db         481 ATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACTTTGGAC 540

Qy         727 AAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCCGAGCTG 786
```

```
Db        541 AAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCCGAGCTG 600

Qy        787 GTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTGCTGCCC 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTGCTGCCC 660

Qy        847 TGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAAATTCCG 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAAATTCCG 720

Qy        907 GAAGGCACCATTCT-AACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAAGCCTA 965
              ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db        721 GAAGGCACCATTCTAAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAAGGCTA 780

Qy        966 TCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAAGAGGA 1025
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        781 TCCTCCTGAAGACATCTTTAATCTAGTGGGCACGAAGGTCCTGGTCACTGTTGAAGAGGA 840

Qy       1026 TGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
              |||||||||||||||||||||||||||||| || ||||||||||||||    | ||
Db        841 TGACACCATCATGGAAGAATTGGTAGATACTCCTGGCAAAAAAATCAGTCTTAAA 895
```

RESULT 9
AL556387

```
LOCUS       AL556387                 987 bp    mRNA    linear   EST 31-MAY-2003
DEFINITION  AL556387 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DK004YI14 5-PRIME, mRNA sequence.
ACCESSION   AL556387
VERSION     AL556387.2  GI:31278189
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 987)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:12899016.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 5047.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DK004BE07QP1&cluster=5047.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DK004BE07QP1.
FEATURES             Location/Qualifiers
     source          1. .987
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DK004YI14"
                     /cell_type="HELA CELLS COT 25-NORMALIZED"
                     /cell_line="HELA"
                     /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
```

```
  Query Match             79.3%;  Score 856.4;  DB 9;  Length 987;
  Best Local Similarity   99.9%;  Pred. No. 5.4e-211;
  Matches  857;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        130 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 189

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        190 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGATGTAGCCAAGCGCATGATC 249

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 309

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 369

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 429

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 489

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 549

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 609

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        610 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 669

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 729

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 789

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 849

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 909

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 969

Qy        841 CTGCCCTGTGAGATGGCC 858
              ||||||||||||||||||
Db        970 CTGCCCTGTGAGATGGCC 987
```

RESULT 10
AL556278

```
LOCUS       AL556278                1005 bp    mRNA    linear   EST 31-MAY-2003
DEFINITION  AL556278 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DK001YM22 5-PRIME, mRNA sequence.
ACCESSION   AL556278
VERSION     AL556278.2  GI:31278081
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1005)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
```

```
 TITLE      Full-length cDNA libraries and normalization
 JOURNAL    Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:12898807.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 5047.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DK001BG11QP1&cluster=5047.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DK001BG11QP1.
FEATURES             Location/Qualifiers
     source          1. .1005
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DK001YM22"
                     /cell_type="HELA CELLS COT 25-NORMALIZED"
                     /cell_line="HELA"
                     /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match            78.1%;  Score 843;  DB 9;  Length 1005;
  Best Local Similarity  95.5%;  Pred. No. 1.7e-207;
  Matches  846;  Conservative    6;  Mismatches   34;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGCGGGCAACACCCGTGCTTC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        120 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGCGGGCAACACCCGTCCTTC 179

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              |||||||||||||||||||||||||||||||||:||||||| ||||||||||||||||:|
Db        180 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCSACCTGGATGTAGCCAAGCGCATGAYC 239

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              |||||||||||||||| | ||||||||| | |||| ||||| || | | | | | |||||
Db        240 CGCATGGCCAAGGAGTNTNGGGCTGATTNTNCCAANTTCCANAANANTNANCCANAATTC 299

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              || | ||||| |||| |||| |||| |||||||||||| || |||||||||||||| |||
Db        300 AANTNTAATCCGAAANCCTTNGAGANGCCATACACCTCCAANCATTCCTGGGGGAANACG 359

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              |||||||||||||||||||||||||||| ||:|||||||||||||||||| || || | |
Db        360 TACGGGGAGCACAAACGACATCTGGAGTCCASCCATGACCAGTACAGGGANCTCCANANG 419

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||| |||| |||||| ||||||||||||||:||||||||||||||||||||||||| |||
Db        420 TACCCCGANGAGGTTNGGATCTTCTTCACTSCCTCTGGCATGGATGAGATGGCAGTCGAA 479

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 539

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 599

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 TYAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 659

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 719

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 779

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 839

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 899

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 959

Qy        841 CTGCCCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGG 886
              |||||||||||||:|||||||||||||||||||||||||| ||| |
Db        960 CTGCCCTGTGAGAWGGCCTGCAATGAGAAGCTGGGCAAGTTTGTTG 1005


RESULT 11
AL515470
LOCUS       AL515470                1201 bp    mRNA    linear   EST 09-MAY-2003
DEFINITION  AL515470 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
            CS0DA002YI24 5-PRIME, mRNA sequence.
ACCESSION   AL515470
VERSION     AL515470.2  GI:30489137
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:12778963.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 5047.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DA002BE12QP1&cluster=5047.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DA002BE12QP1.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DA002YI24"
                     /tissue_type="NEUROBLASTOMA"
                     /clone_lib="Homo sapiens NEUROBLASTOMA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

  Query Match            77.8%;  Score 840;  DB 9;  Length 1201;
  Best Local Similarity  98.7%;  Pred. No. 1.1e-206;
  Matches  859;  Conservative    7;  Mismatches    2;  Indels    2;  Gaps    2;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGCGGGCAACACCCGTGCTTC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGCGGGCAACACCCGTGCTTC 162
```

```
Qy        61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
             ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       163 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGATGTAGCCAAGCGCATGATC 222

Qy       121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       223 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 282

Qy       181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       283 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 342

Qy       241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       343 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 402

Qy       301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       403 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 462

Qy       361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       463 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 522

Qy       421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       523 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 582

Qy       481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       583 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 642

Qy       541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       643 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 702

Qy       601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       703 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 762

Qy       661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
             ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db       763 ATAGCGATATCTGTKGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 822

Qy       721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
             |||||||||| |||||||||||:|||||||||||||||||| :|||||||||||||||||
Db       823 TTGGACAAGA-CTGGAAGGGGAGTKACCACTCGGCCTCGCTGGAGCTKGGAGAACTGGCC 881

Qy       781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTG 840
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||:|
Db       882 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCYGGGCTCCCCAACCAAGCAGCYG 941

Qy       841 CTGCCCTGTGAGATGGCCTGCAATGAGAAG 870
             |||||||||||||:|| ||||||||||:|
Db       942 CTGCCCTGTGAGAKGG-CTGCAATGAGARG 970
```

```
RESULT 12
AL526241/c
LOCUS       AL526241               989 bp    mRNA    linear   EST 23-MAY-2003
DEFINITION  AL526241 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DC016YM22 3-PRIME, mRNA sequence.
ACCESSION   AL526241
VERSION     AL526241.2  GI:31064102
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 989)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:12789734.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 5047.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DC016BG11NP1&cluster=5047.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DC016BG11NP1.
FEATURES              Location/Qualifiers
     source           1. .989
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DC016YM22"
                      /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
                      /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
                      /note="1st strand cDNA was primed with a NotI-oligo(dT)
                      primer. Five prime end enriched, double-strand cDNA was
                      digested with Not I and cloned into the Not I and EcoR V
                      sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
```

Query Match 76.9%; Score 830; DB 9; Length 989;
Best Local Similarity 93.0%; Pred. No. 3.9e-204;
Matches 872; Conservative 34; Mismatches 26; Indels 6; Gaps 4;

```
Qy       126 GGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTCAAGTT 185
             :|||||||||||||||      ||    | |  |||||||||:||||||||||||||||||
Db       940 KGCCAAGGAGTGTGGGCTGATGTG---CTAAGTCCAGAAGAKTGAGCTAGAATTCAAGTT 884

Qy       186 TAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACGTACGG 245
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       883 TAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACGTACGG 824

Qy       246 GGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGGTACGC 305
             |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db       823 GGAGCACAAACGACWTCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGGTACGC 764

Qy       306 CGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAATTCCT 365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       763 CGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAATTCCT 704

Qy       366 GCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTTCCTTA 425
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db       703 GCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTTCCTTW 644

Qy       426 TCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAGTCAAT 485
             |:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       643 TYTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAGTCAAT 584

Qy       486 GGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGCTTCTT 545
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       583 GGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGCTTCTT 524

Qy       546 GCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATCTCGGA 605
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       523 GCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAA-CTGCGGGTCATCTCGGA 465

Qy       606 ATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGCATAGC 665
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       464 ATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAGACAGGCATAGC 405

Qy       666 GATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACTTTGGA 725
             ||||||||||||||||:||| |:|||::|||||||||||||||||||||||||||||||||
Db       404 GATATCTGTGGCCGCAGKGGCCCWGGGRRCCAAGGTGTTGGAACGTCACATAACTTTGGA 345

Qy       726 CAAGACCTGGAAGGGGAGTG-ACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCCGAGC 784
```

```
Db        344 CAAGACCTGGAAGGGGAGTGCACCACCCSVCCTCSCHGGAGCCTGGAGNACTGGCCGAGC 285

Qy        785 TGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTGCTGC 844

Db        284 TGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCCMMCCAAGCAGCTGCTGC 225

Qy        845 CCTGTGAGATGGCCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAAATTC 904

Db        224 CCTGTGAGATGGCCTSCAATGAGAAGCTGSGCAAGTCTGTGGTGGCCAAAGTGAAAATTC 165

Qy        905 CGGAAGGCACCATTCTAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAAGCCT 964

Db        164 CGGAAGGCACCABTCTMACAATGCACATGCYCACCGTGAAGGTGGGTGAMCCCAAAGGCT 105

Qy        965 ATCCTCCTGAAGACATCTTTAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAAGAGG 1024

Db        104 AGCCGCCCGAAGACVMCTTTAATCTANTGGGCAAGAAGGYCCTGGWVACTGTTGAAGAGG 45

Qy       1025 ATGACACC-ATCATGGAAGAATTGGTAGATAATCATGG 1061

Db         44 ATGACACCTNTNATGGGAGWATTGGKDGDKRRKTSDKG 7
```

RESULT 13
BU187658

LOCUS	BU187658 889 bp mRNA linear EST 04-SEP-2002
DEFINITION	AGENCOURT_7851896 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6144738 5', mRNA sequence.
ACCESSION	BU187658
VERSION	BU187658.1 GI:22701642
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 889)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM13470 row: d column: 19 High quality sequence stop: 724.

FEATURES	Location/Qualifiers
source	1. .889 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6144738" /tissue_type="retinoblastoma" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_67" /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."

ORIGIN

Query Match 76.2%; Score 823.4; DB 13; Length 889;
Best Local Similarity 99.1%; Pred. No. 1.9e-202;
Matches 838; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

```
Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60

Db         44 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 103

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120

Db        104 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 163

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180

Db        164 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCTAAGTTCCAGAAGAGTGAGCTAGAATTC 223

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240

Db        224 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 283

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300

Db        284 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 343

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360

Db        344 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 403

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420

Db        404 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 463

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480

Db        464 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 523

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540

Db        524 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 583

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600

Db        584 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 643

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660

Db        644 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 703

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720

Db        704 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 763

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780

Db        764 TTGGACAAGACCTGGGAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 823

Qy        781 GAGCTGGTGCGGTCAGTGCG-TCTTGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCT 839

Db        824 GAGCTGTTGCGNTCAGTGCGTTCTTGTGGAGCGTGCCCTGGGGCTCCCAACCAAGCAGCT 883

Qy        840 GCTGCC 845

Db        884 GCTGCC 889
```

RESULT 14
BQ432284

LOCUS	BQ432284 872 bp mRNA linear EST 24-MAY-2002
DEFINITION	AGENCOURT_7858941 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6167793 5', mRNA sequence.
ACCESSION	BQ432284
VERSION	BQ432284.1 GI:21171360
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 872)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)

```
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13530  row: e  column: 10
            High quality sequence stop: 711.
FEATURES             Location/Qualifiers
     source          1. .872
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6167793"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             74.4%;  Score 803.2;  DB 13;  Length 872;
  Best Local Similarity   99.0%;  Pred. No. 3.3e-197;
  Matches  808;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         57 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGCGGGCAACACCCGTGCTTC 116

Qy         61 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 ATCATTGCCGAGATCGGCCAGAACCACCAGGGCGACCTGGACGTAGCCAAGCGCATGATC 176

Qy        121 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAATTC 180
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        177 CGCATGGCCAAGGAGTGTGGGGCTGATTGTGCTAAGTTCCAGAAGAGTGAGCTAGAATTC 236

Qy        181 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 AAGTTTAATCGGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 296

Qy        241 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 TACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGG 356

Qy        301 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAA 416

Qy        361 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 TTCCTGCATGAACTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 476

Qy        421 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 CCTTATCTGGAAAAGACAGCCAAAAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAG 536

Qy        481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGC 596

Qy        541 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 TTCTTGCAGTGTACCAGCGCATACCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATC 656

Qy        601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        657 TCGGAATATCAGAAGCTCTTTCCTGACATTCCCATAGGGTATTCTGGGCATGAAACAGGC 716

Qy        661 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        717 ATAGCGATATCTGTGGCCGCAGTGGCTCTGGGGGCCCAGGTGTTGGAACGTCACATAACT 776

Qy        721 TTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCC 780
              |||||||||||  | ||||||||||||||||||||||||||||  ||||||||||||||
Db        777 TTGGACAAGACTGGAAAGGGGAGTGACCACTCGGCCTCGCTGGGACCTGGAGAACTGGCC 836

Qy        781 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAGCGTGCC 816
              |||||||||||||||||||||||||||||| ||||||
Db        837 GAGCTGGTGCGGTCAGTGCGTCTTGTGGAACGTGCC 872


RESULT 15
BQ064521
LOCUS       BQ064521                1031 bp    mRNA    linear   EST 02-APR-2002
DEFINITION  AGENCOURT_6853542 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5926258
            5', mRNA sequence.
ACCESSION   BQ064521
VERSION     BQ064521.1  GI:19893231
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1031)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2099  row: e  column: 11
            High quality sequence start: 22
            High quality sequence stop: 706.
FEATURES             Location/Qualifiers
     source          1. .1031
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5926258"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_99"
                     /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match             73.7%;  Score 796.4;  DB 13;  Length 1031;
  Best Local Similarity   99.1%;  Pred. No. 2.1e-195;
  Matches  811;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy        264 GGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGGTACGCCGAGGAGGTTGGGATCTT 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 GGAGTTCAGCCATGACCAGTACAGGGAGCTGCAGAGGTACGCCGAGGAGGTTGGGATCTT 82

Qy        324 CTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAATTCCTGCATGAACTGAATGTTCC 383
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        83 CTTCACTGCCTCTGGCATGGATGAGATGGCAGTTGAATTCCTGCATGAACTGAATGTTCC 142

Qy       384 ATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTTCCTTATCTGGAAAAGACAGCCAA 443
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       143 ATTTTTCAAAGTTGGATCTGGAGACACTAATAATTTTCCTTATCTGGAAAAGACAGCCAA 202

Qy       444 AAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAGTCAATGGACACCATGAAGCAAGT 503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       203 AAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGCAGTCAATGGACACCATGAAGCAAGT 262

Qy       504 TTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGCTTCTTGCAGTGTACCAGCGCATA 563
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       263 TTATCAGATCGTGAAGCCCCTCAACCCCAACTTCTGCTTCTTGCAGTGTACCAGCGCATA 322

Qy       564 CCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATCTCGGAATATCAGAAGCTCTTTCC 623
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       323 CCCGCTCCAGCCTGAGGACGTCAACCTGCGGGTCATCTCGGAATATCAGAAGCTCTTTCC 382

Qy       624 TGACATTCCCATAGGGTATTCTGGGCATGAAACAGGCATAGCGATATCTGTGGCCGCAGT 683
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       383 TGACATTCCCATAGGGTATTCTGGGCATGAAACAGGCATAGCGATATCTGTGGCCGCAGT 442

Qy       684 GGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACTTTGGACAAGACCTGGAAGGGGAG 743
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       443 GGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACTTTGGACAAGACCTGGAAGGGGAG 502

Qy       744 TGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCT 803
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       503 TGACCACTCGGCCTCGCTGGAGCCTGGAGAACTGGCCGAGCTGGTGCGGTCAGTGCGTCT 562

Qy       804 TGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAA 863
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       563 TGTGGAGCGTGCCCTGGGCTCCCCAACCAAGCAGCTGCTGCCCTGTGAGATGGCCTGCAA 622

Qy       864 TGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAAC 923
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       623 TGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAAATTCCGGAAGGCACCATTCTAAC 682

Qy       924 AATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAAGCCTATCCTCCTGAAGACATCTT 983
             |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db       683 AATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAAAGCTATCCTCCTGAAGACATCTT 742

Qy       984 TAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAAGAGGATGACACCATCATGG-AAG 1042
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       743 TAATCTAGTGGGCAAGAAGGTCCTGGTCACTGTTGAAGAGGATGACACCATCATGGAAAG 802

Qy      1043 AATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
             |||||||| ||||||||||||||||||   ||||||||
Db       803 AATTGGTAAATAATCATGGCAAAAAAAATCAGTCTTAA 840
```

Search completed: September 16, 2004, 05:19:48
Job time : 4423 secs